

Characterization of missingness in untargeted MS-based metabolomics

data sets and evaluation of missing data handling strategies

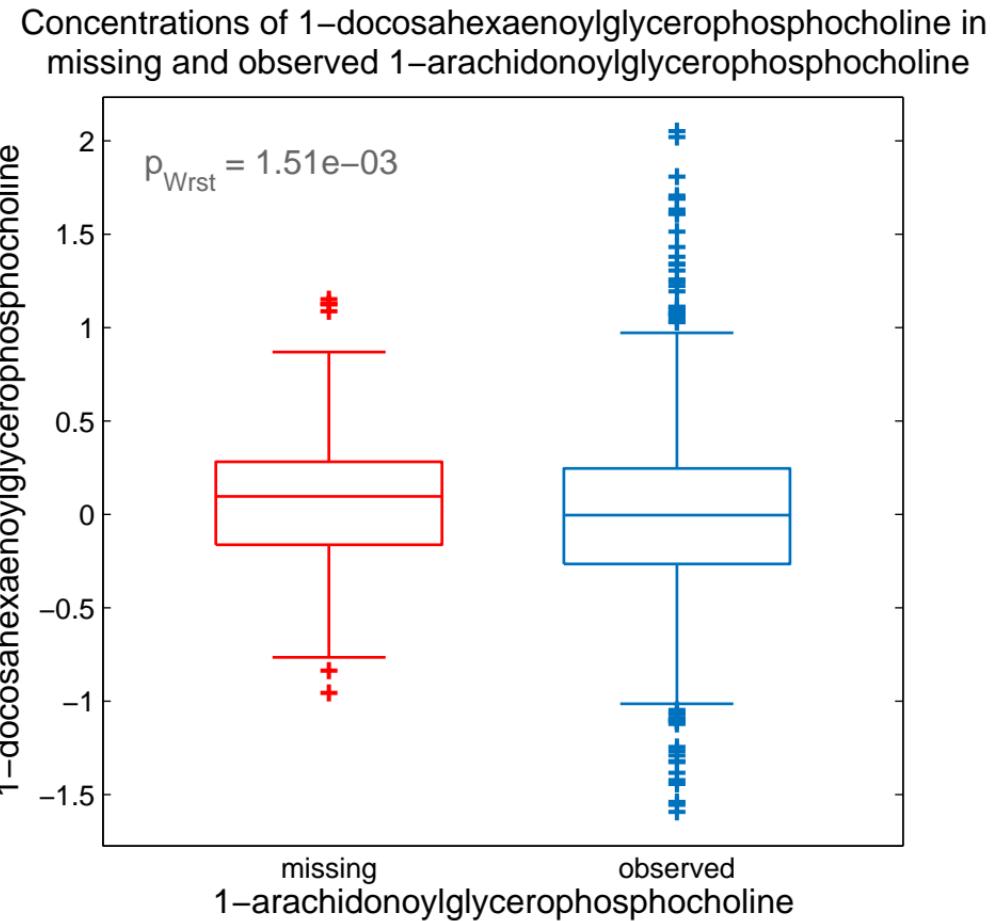
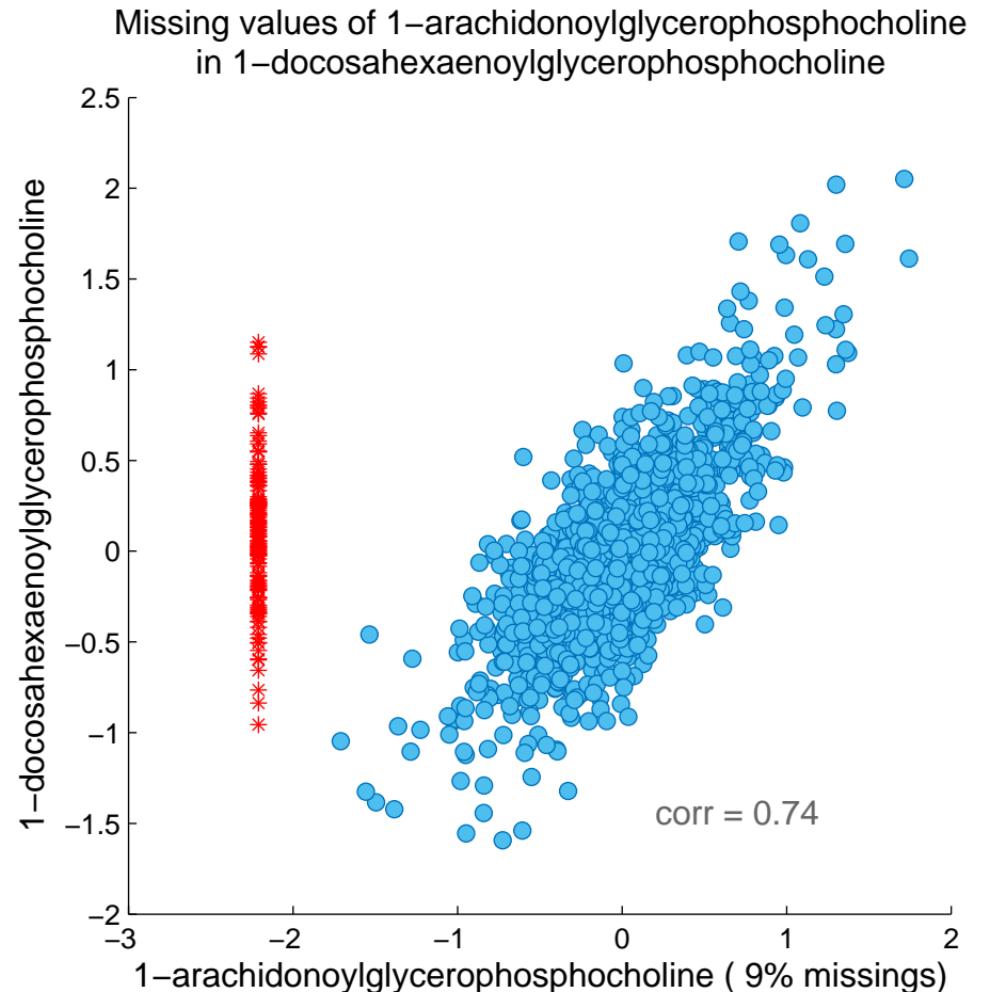
Kieu Trinh Do[†], Simone Wahl[†], Johannes Raffler, Sophie Molnos, Michael Laimighofer, Jerzy Adamski, Karsten Suhre, Konstantin Strauch, Annette Peters, Christian Gieger, Claudia Langenberg, Isobel D. Stewart, Fabian J. Theis, Harald Grallert, Gabi Kastenmüller[#], Jan Krumsiek[#]

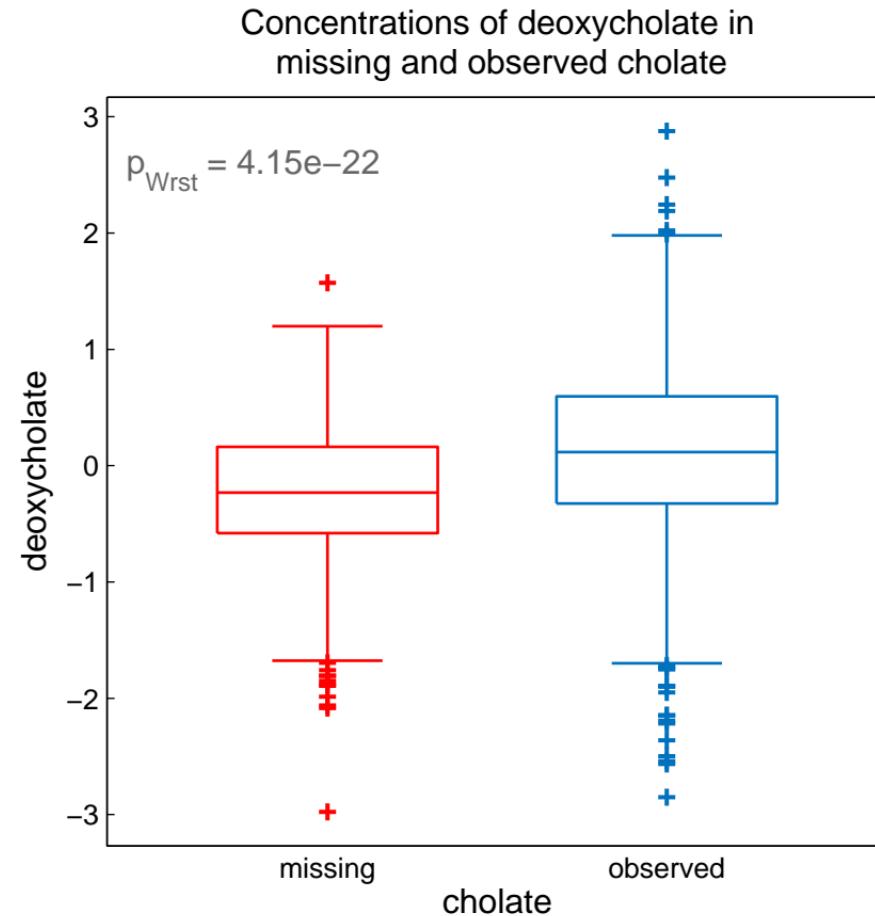
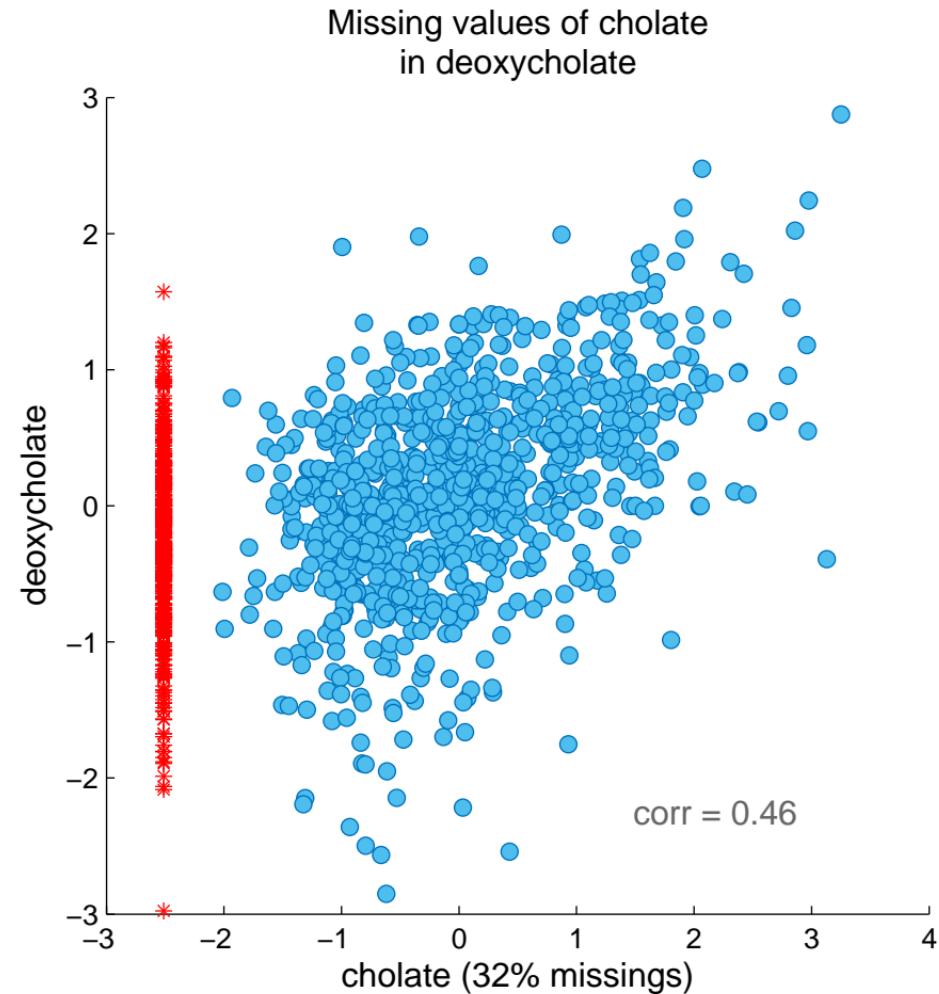
Supporting Information File S1: LOD tendency

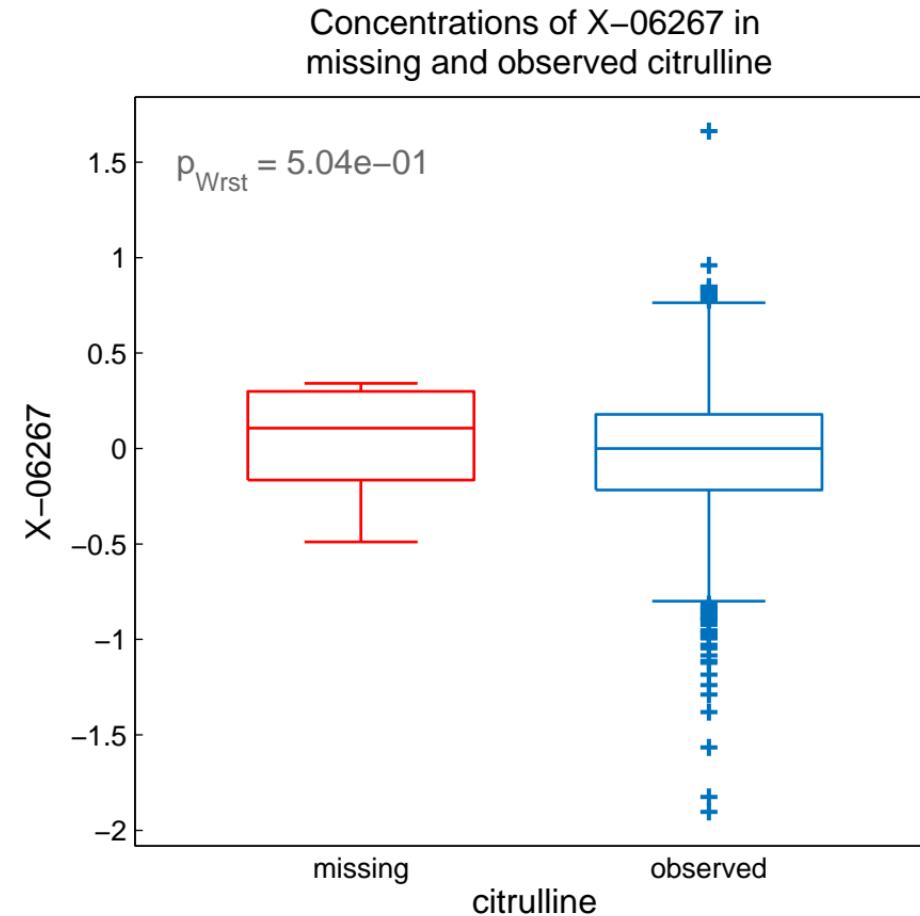
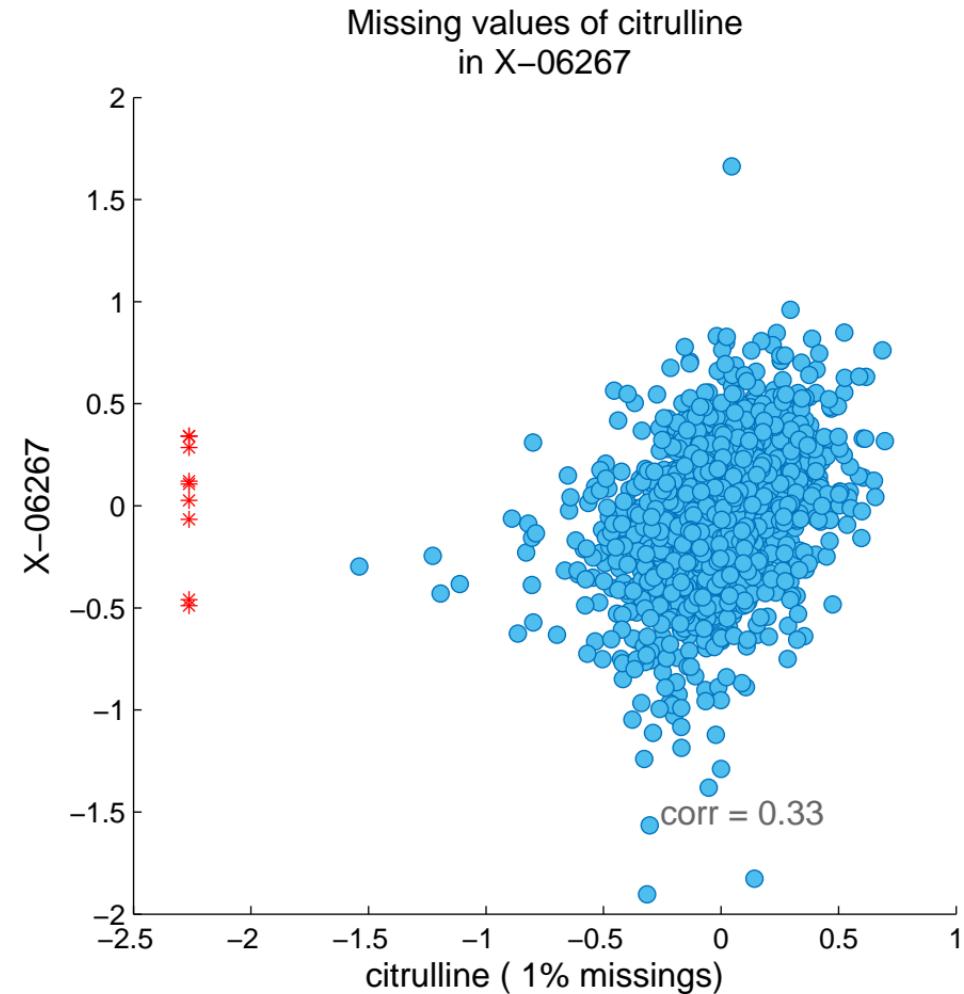
As an explorative method to visualize the missingness mechanism of a metabolite m , a second (auxiliary) metabolite m_{aux} showing the strongest Pearson correlation to m was determined. Comparing the distribution of concentrations of metabolite m_{aux} for observations with and without observed m can provide information about whether or not missingness in m is systematic.

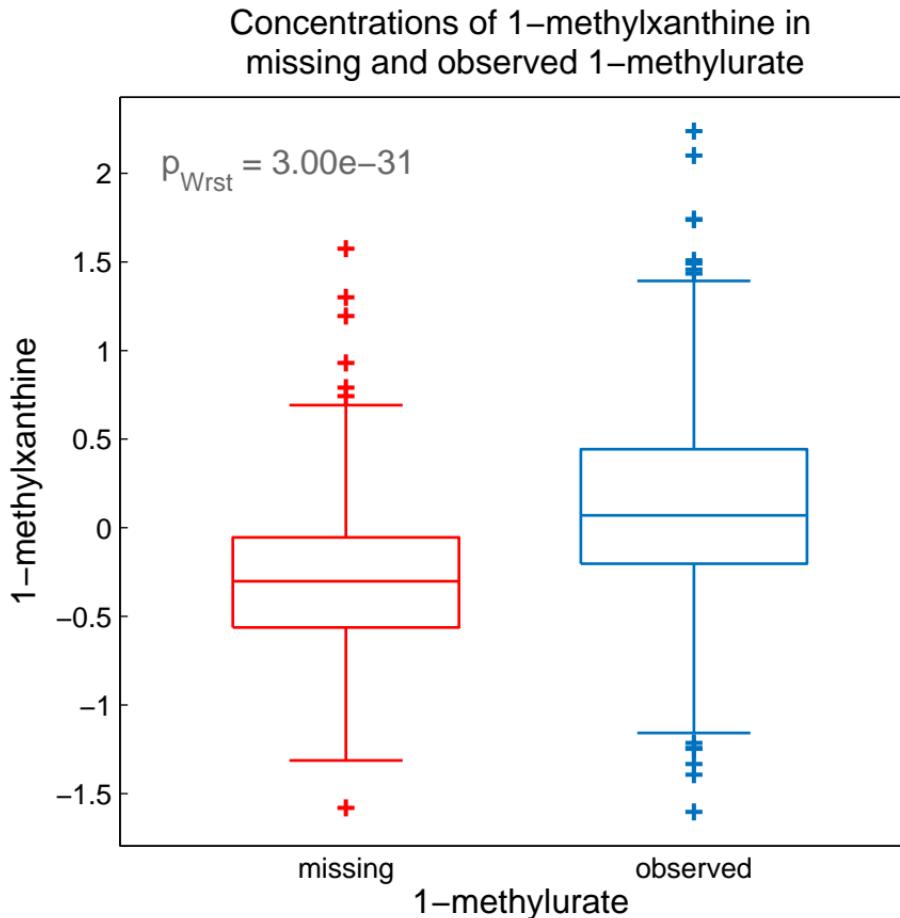
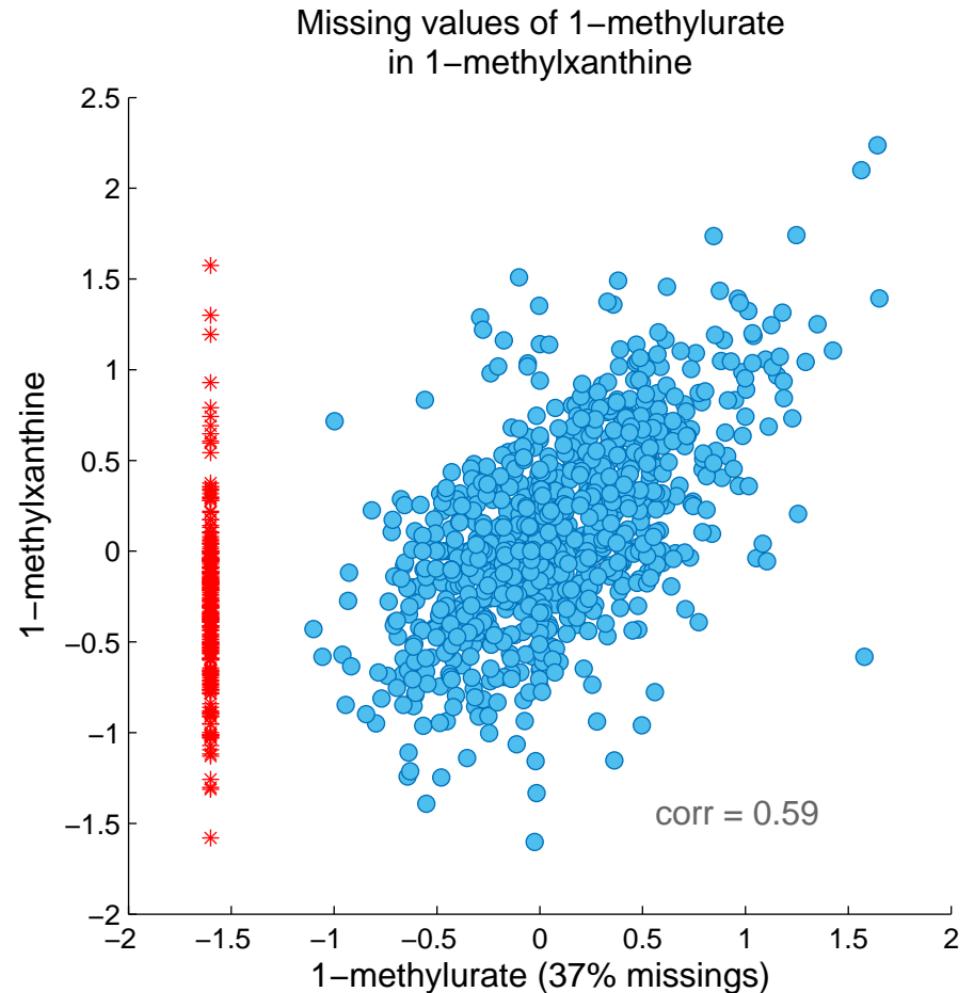
In the following, for all metabolites (x-axis) with more than 10% and less than 70% missing values a scatterplot with its auxiliary metabolite (y-axis) is shown. Blue dots correspond to observed concentrations in both metabolites. Red dots represent concentrations of the auxiliary metabolite in samples with missing values of the investigated metabolite. **corr** and **p** are Pearson correlation coefficient and corresponding p-value.

The boxplots compare the concentrations of the auxiliary metabolite between samples with missing and samples with measured values of the given metabolite. **p_{Wst}** is the p-value of a performed Wilcoxon-Mann-Whitney test. An LOD tendency is assumed if **p_{Wst}** is below $\alpha = 0.05$ after Bonferroni correction for multiple testing.

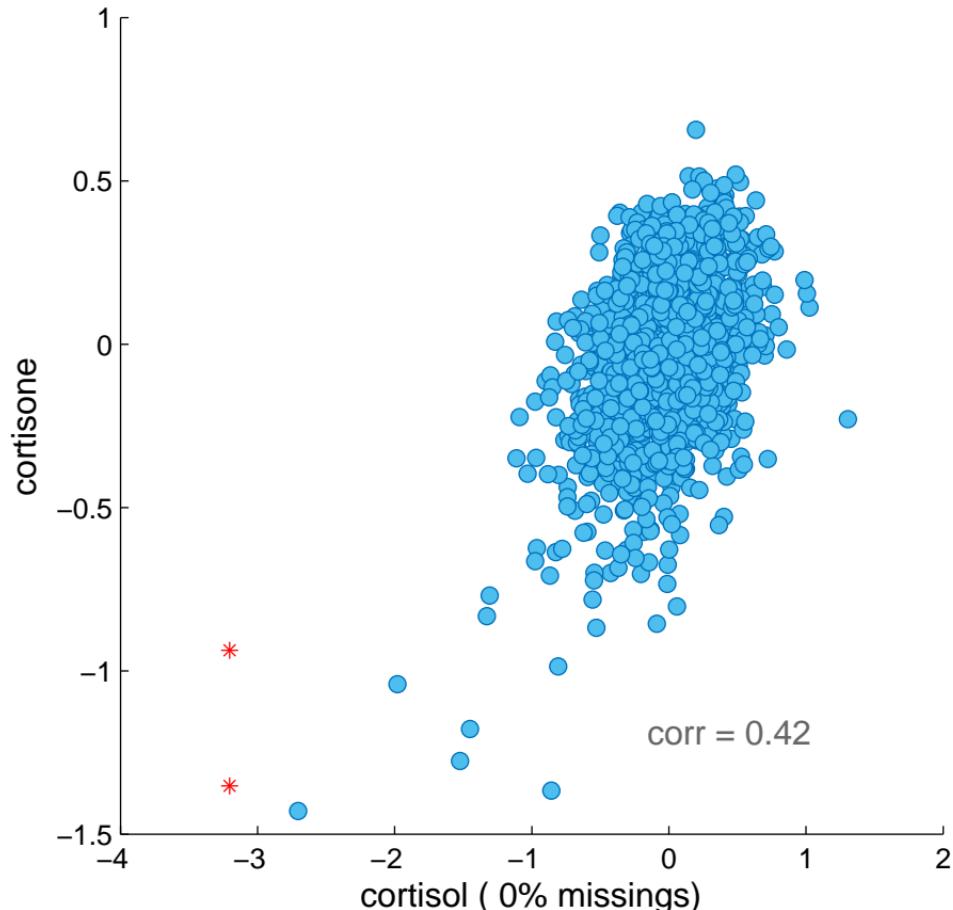




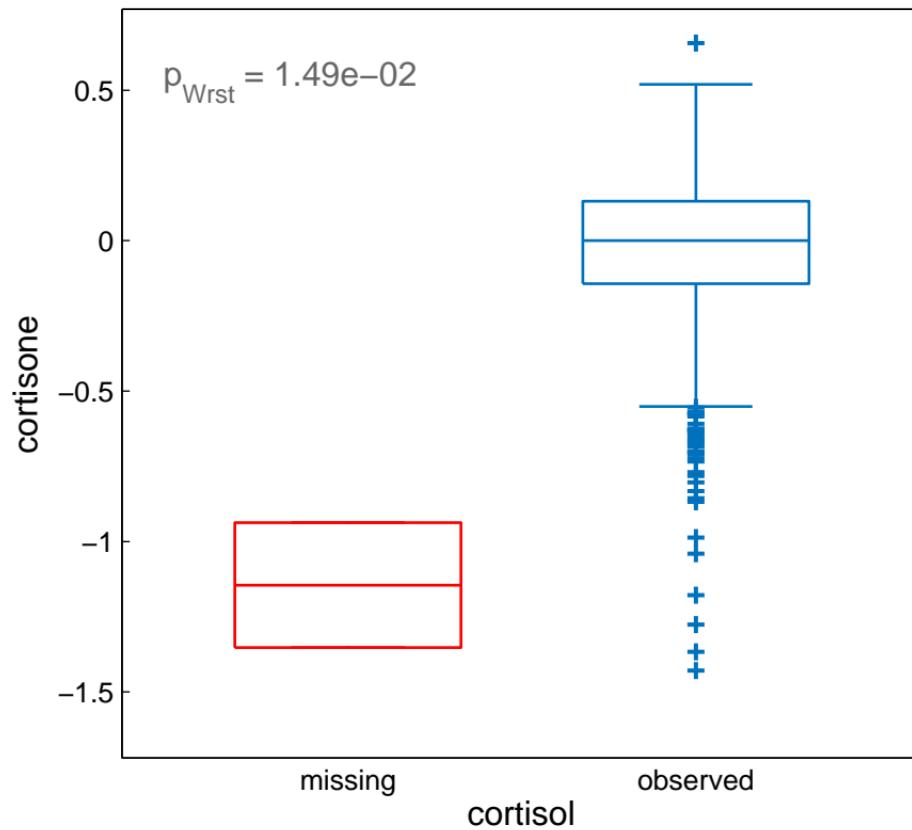


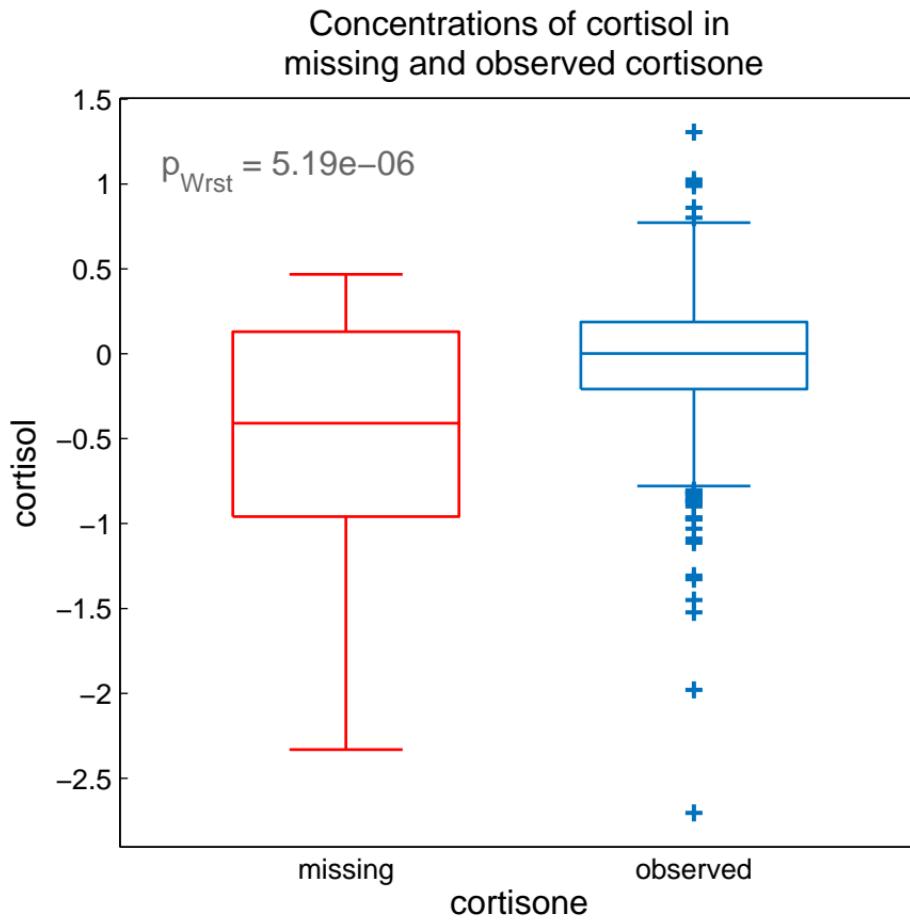
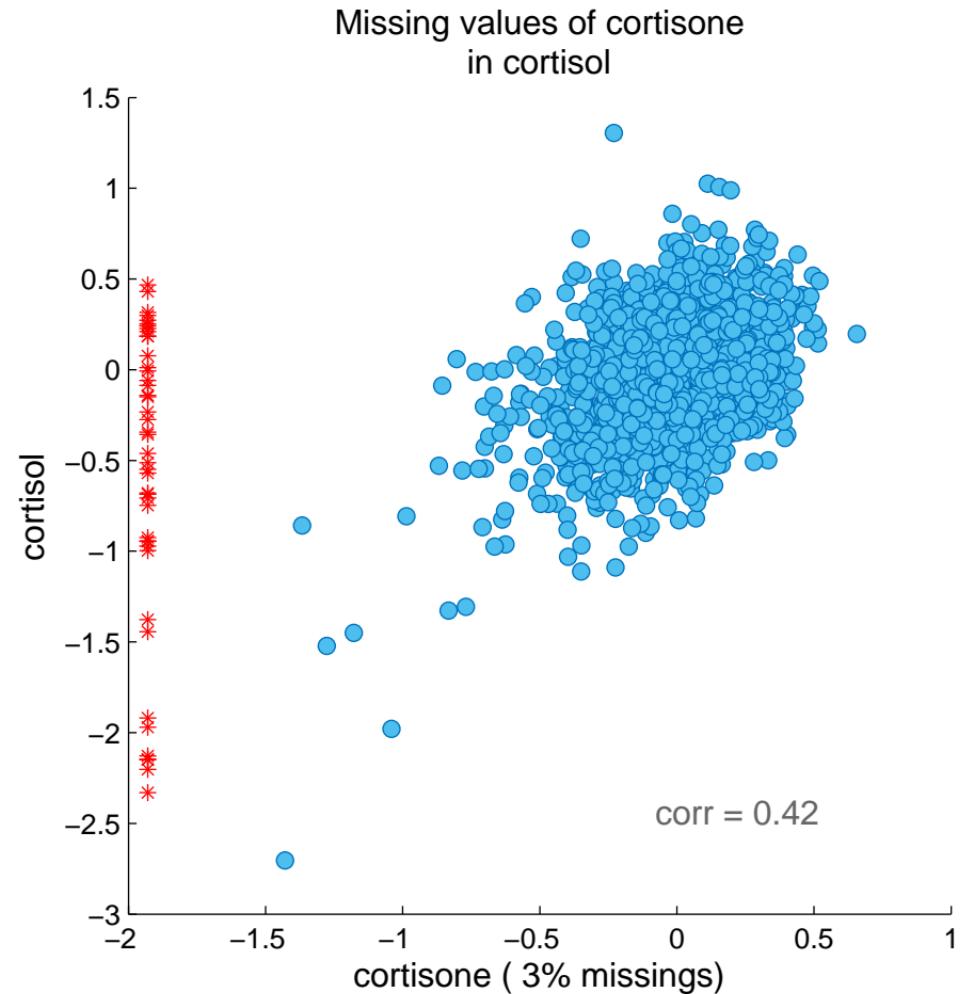


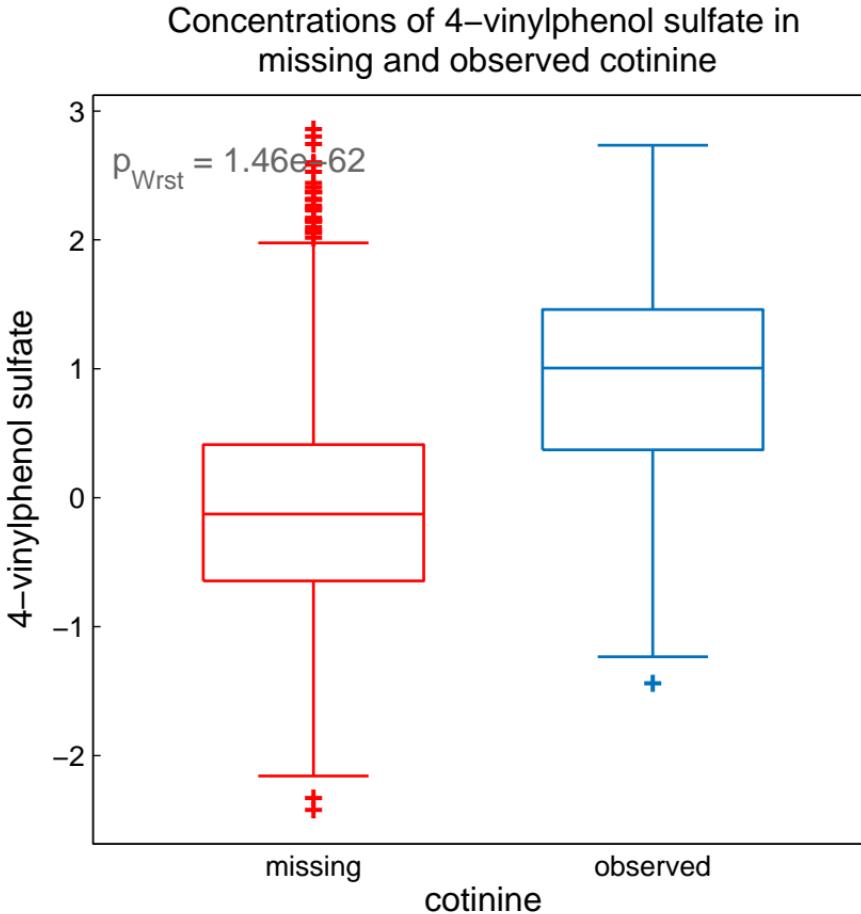
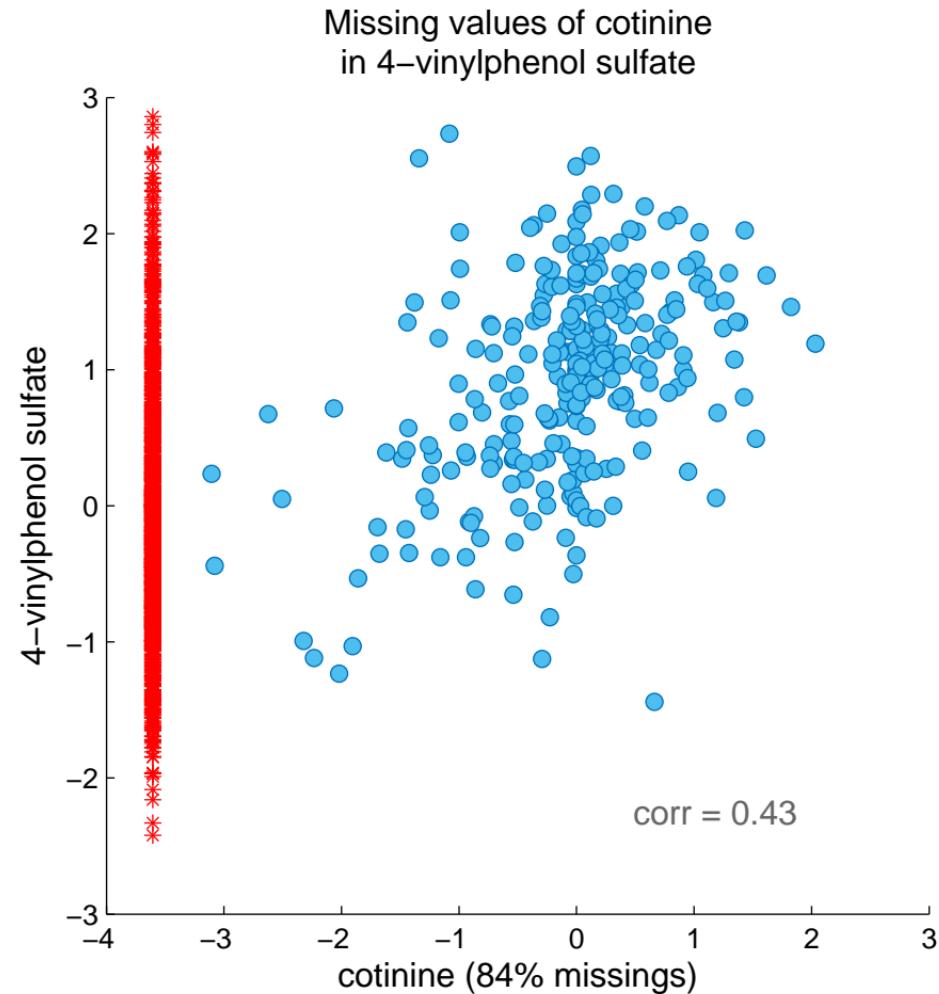
Missing values of cortisol
in cortisone

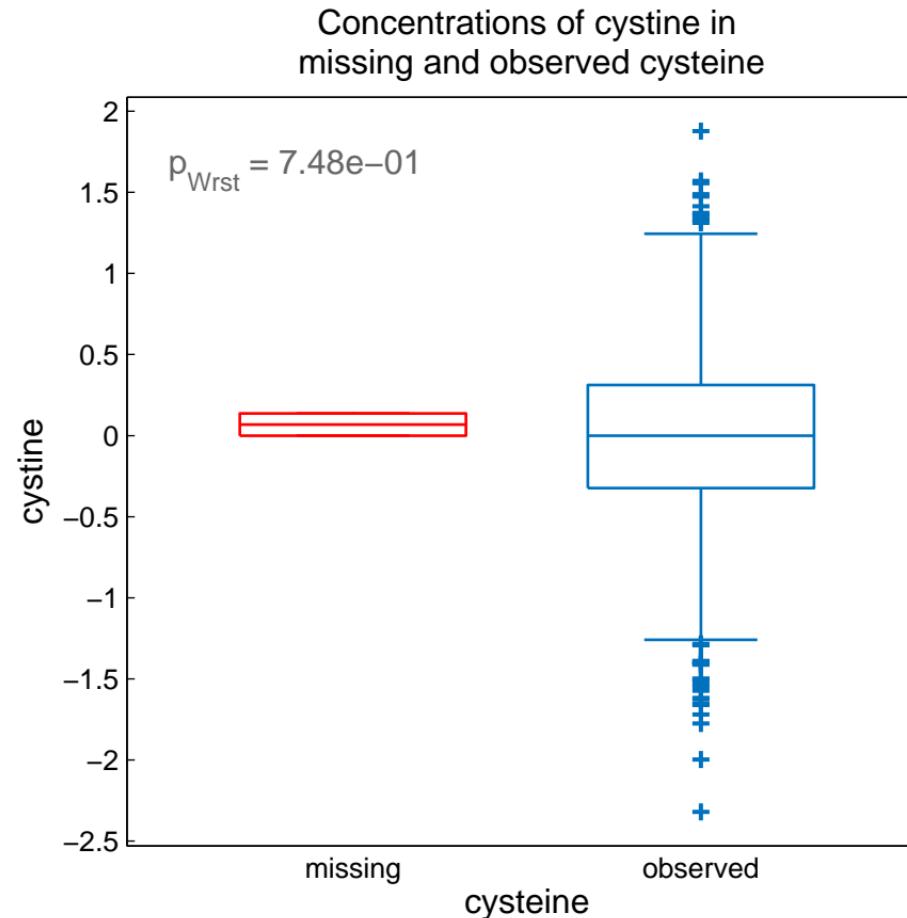
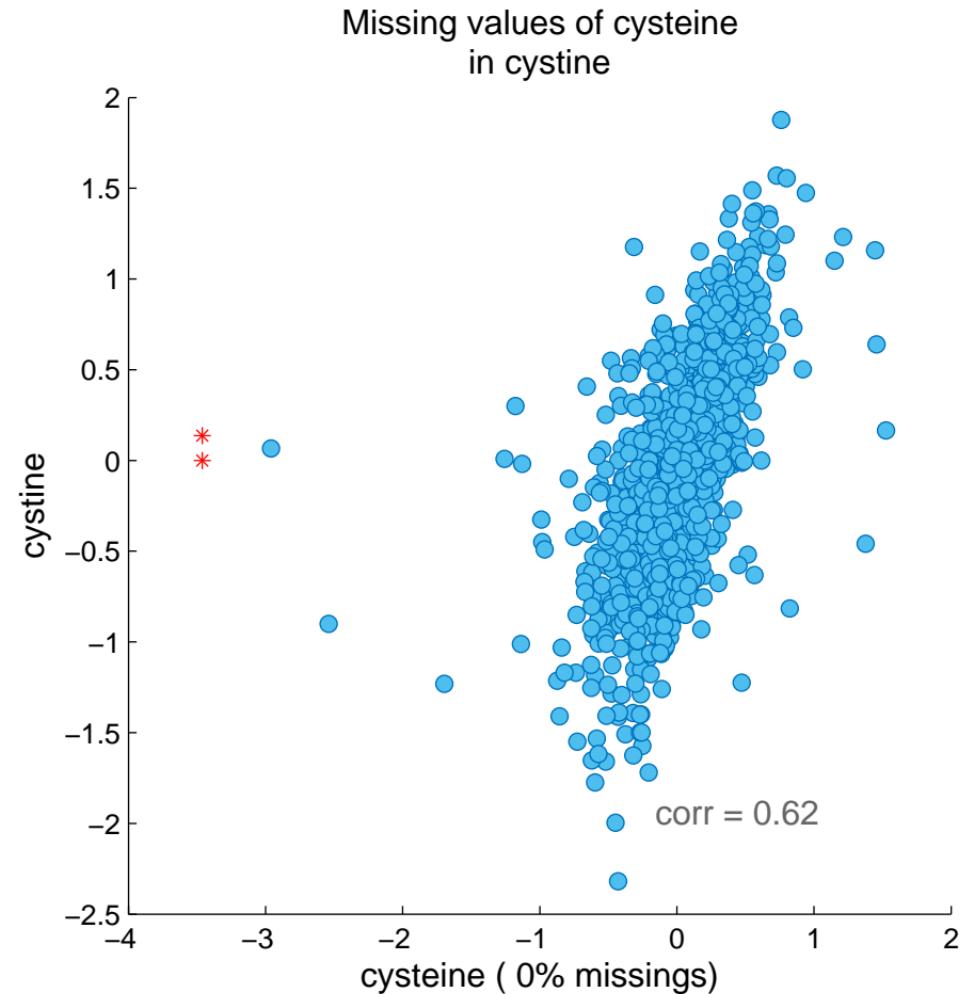


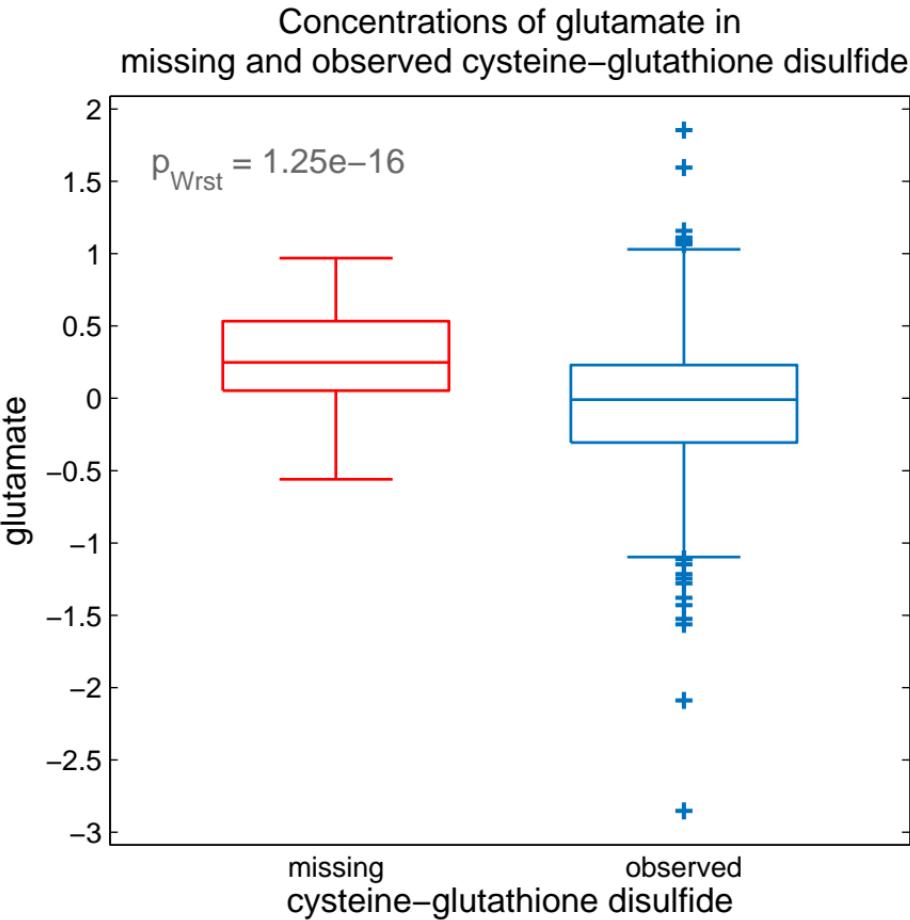
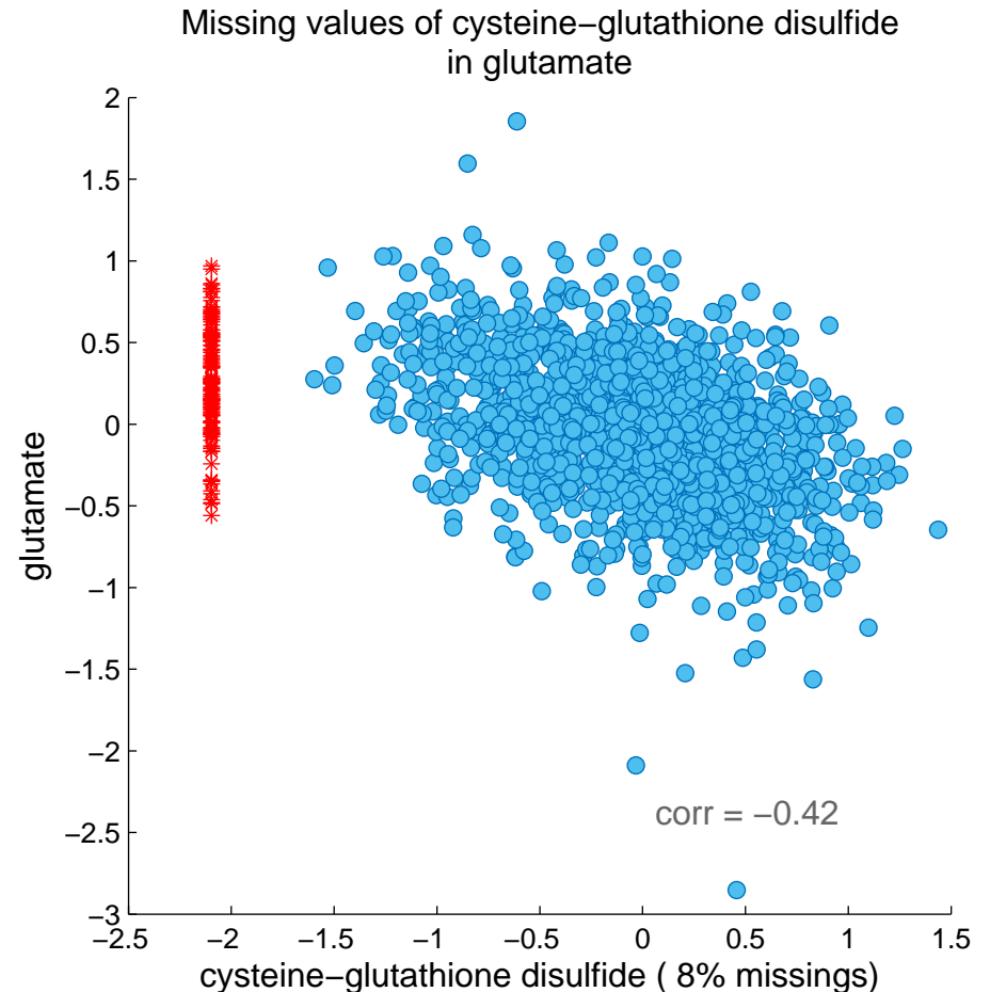
Concentrations of cortisone in
missing and observed cortisol

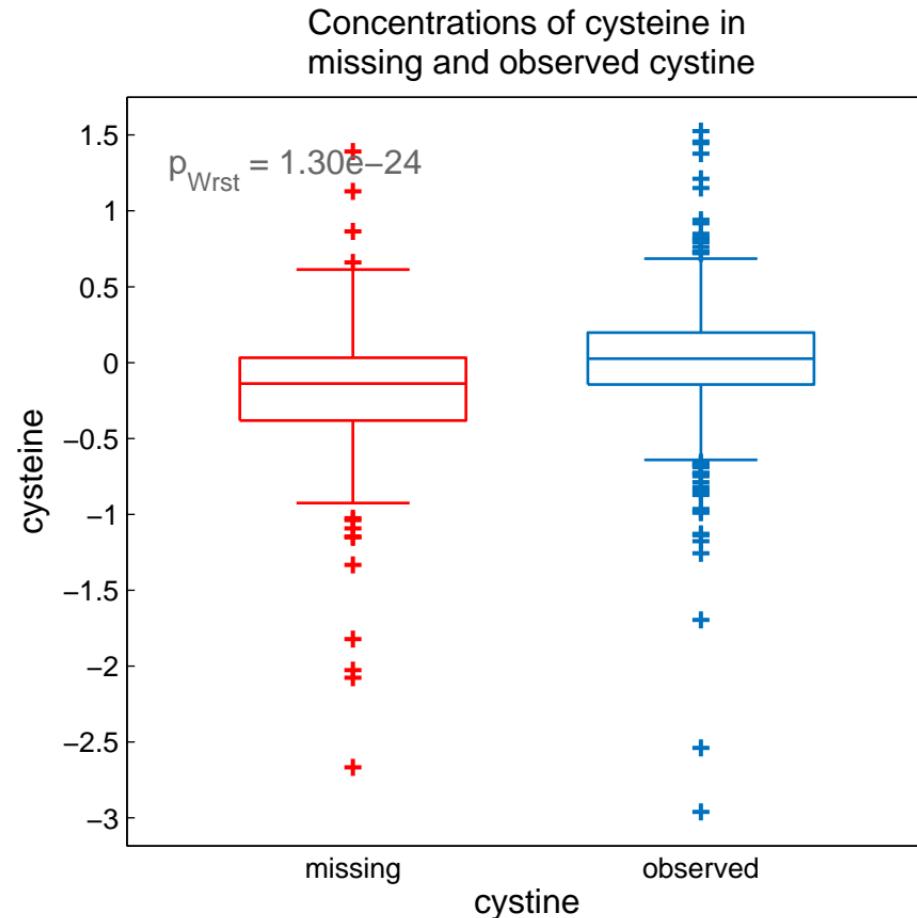
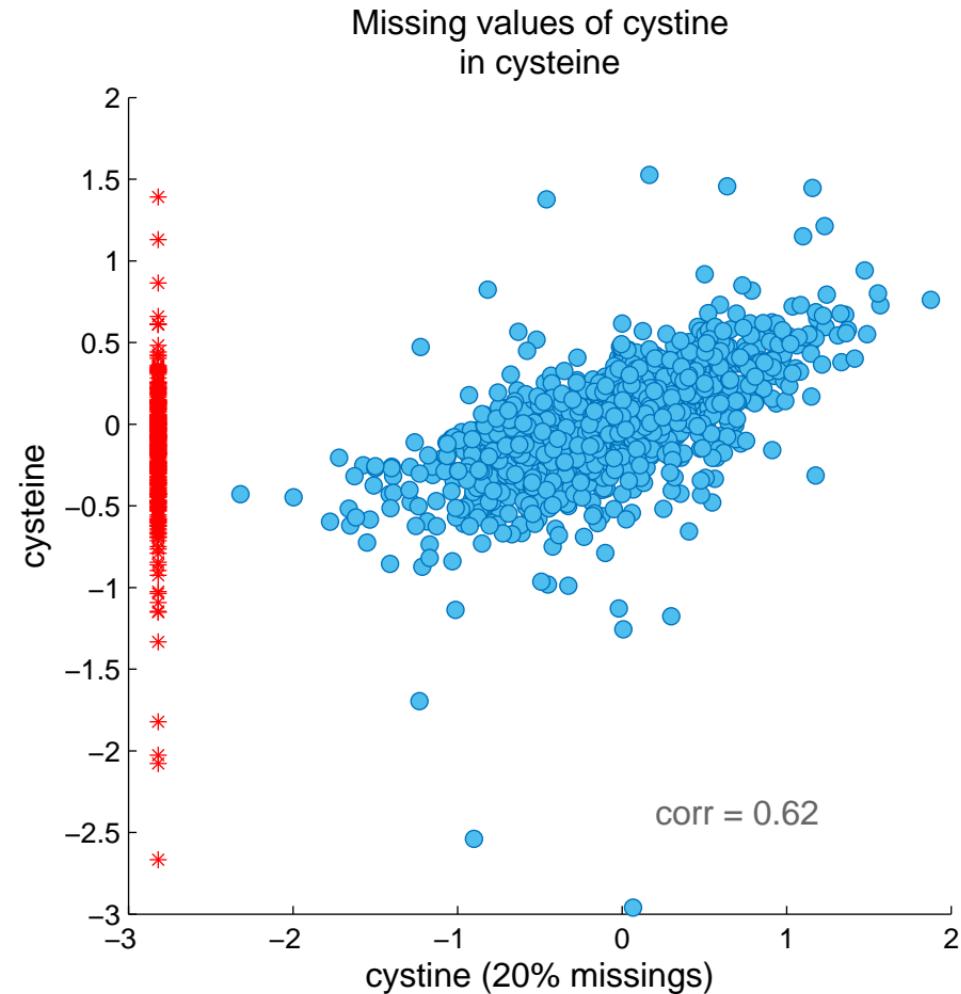


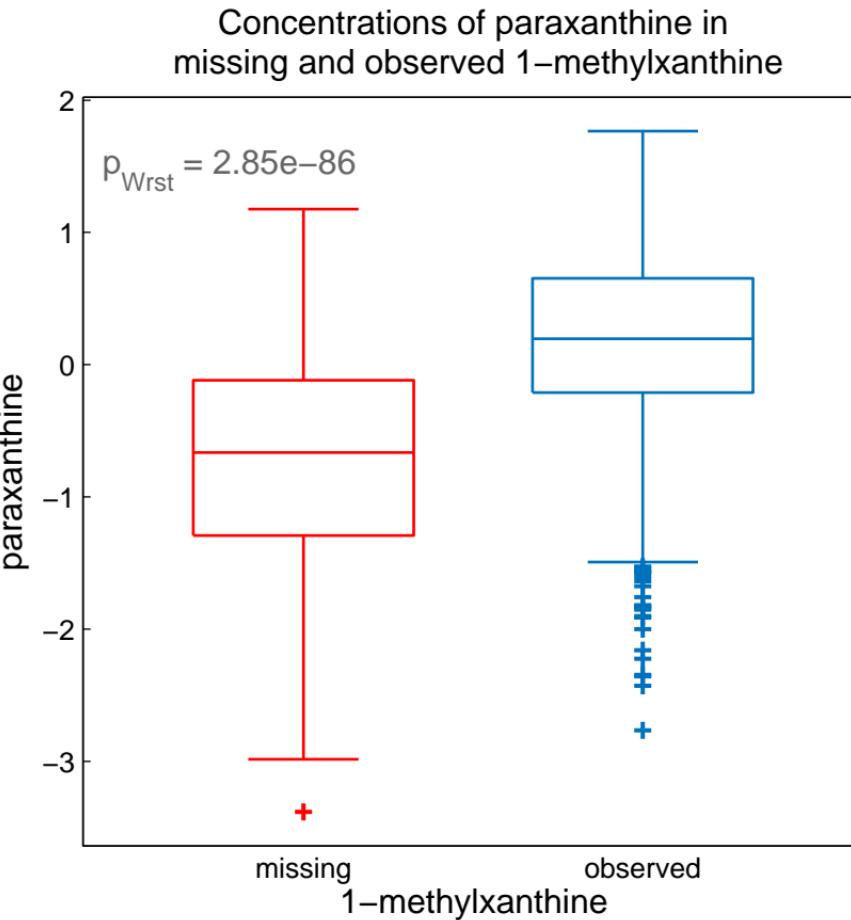
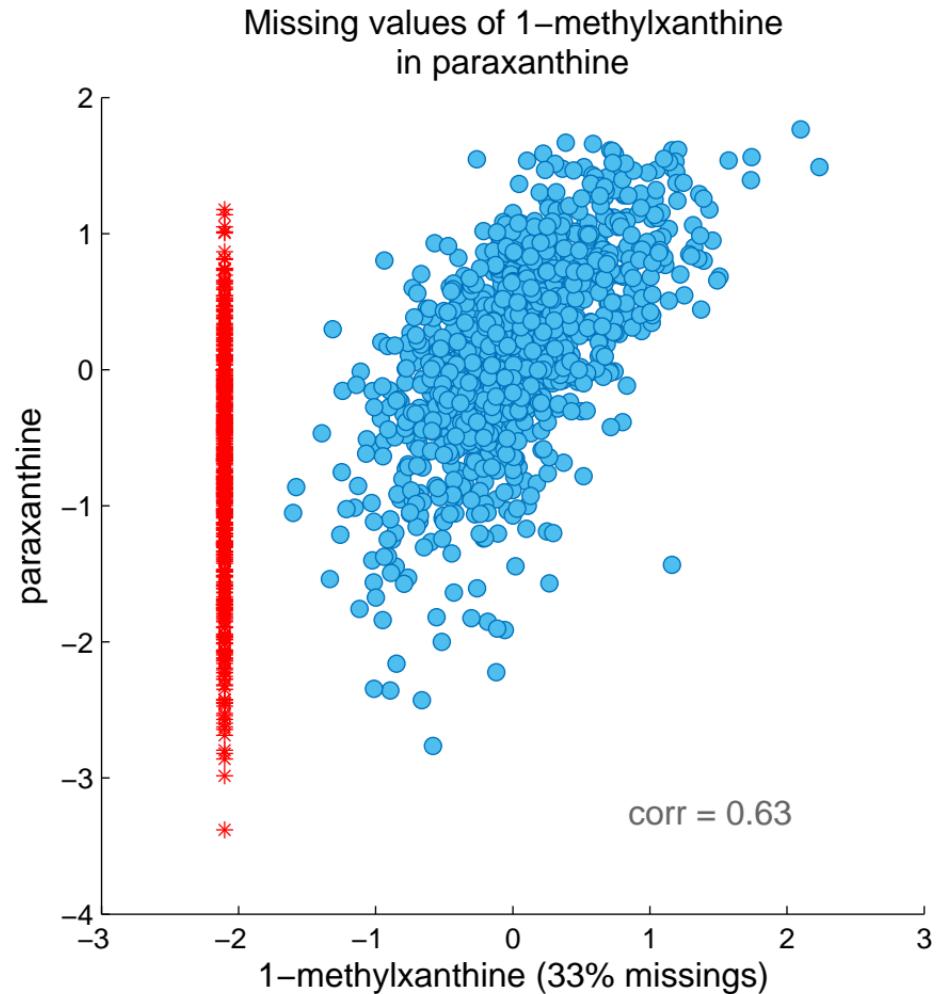


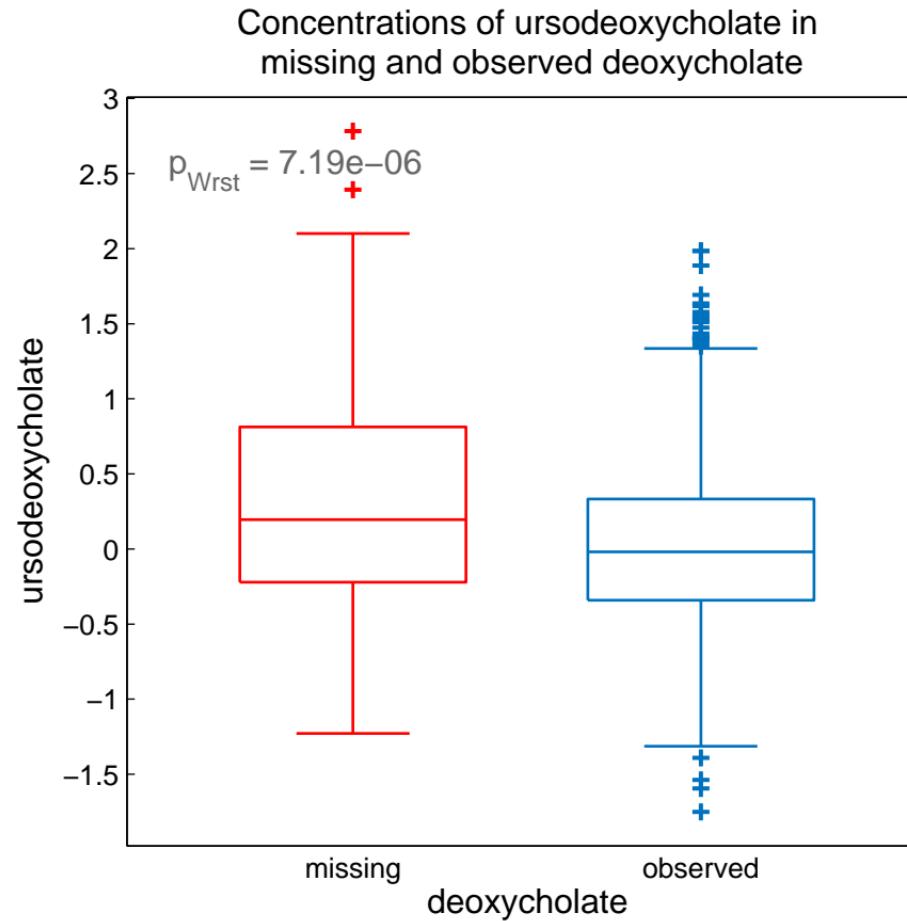
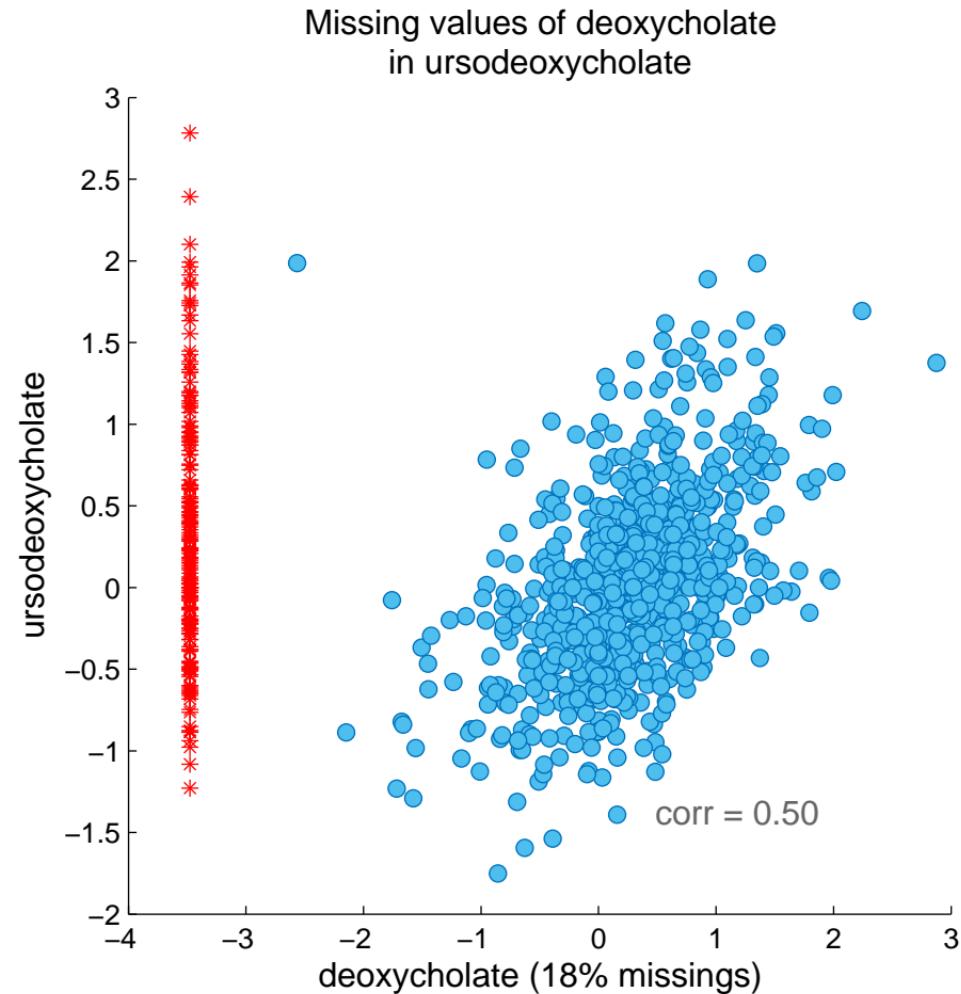


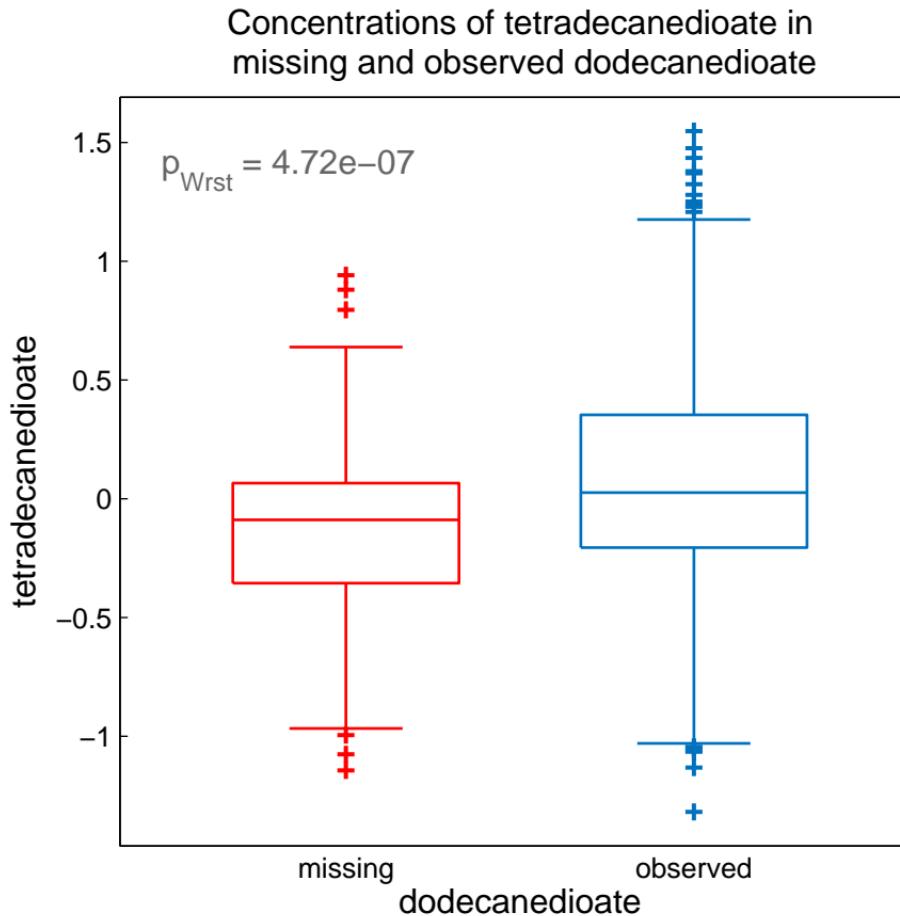
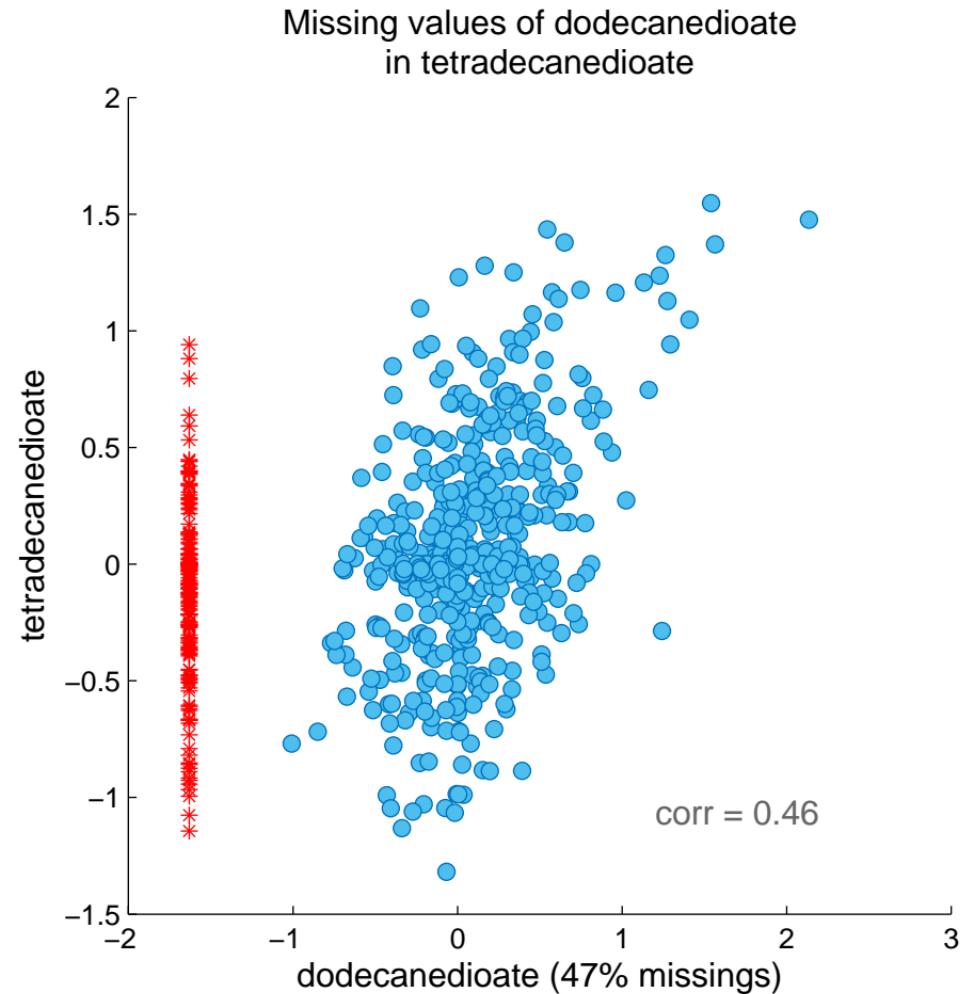


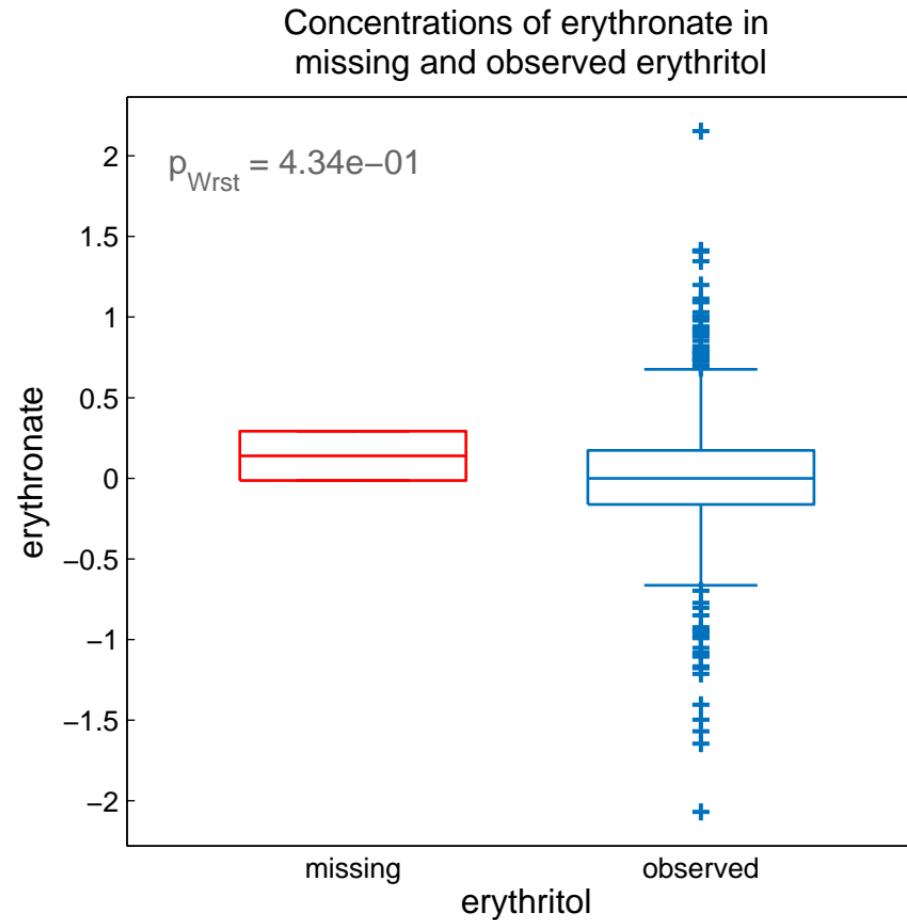
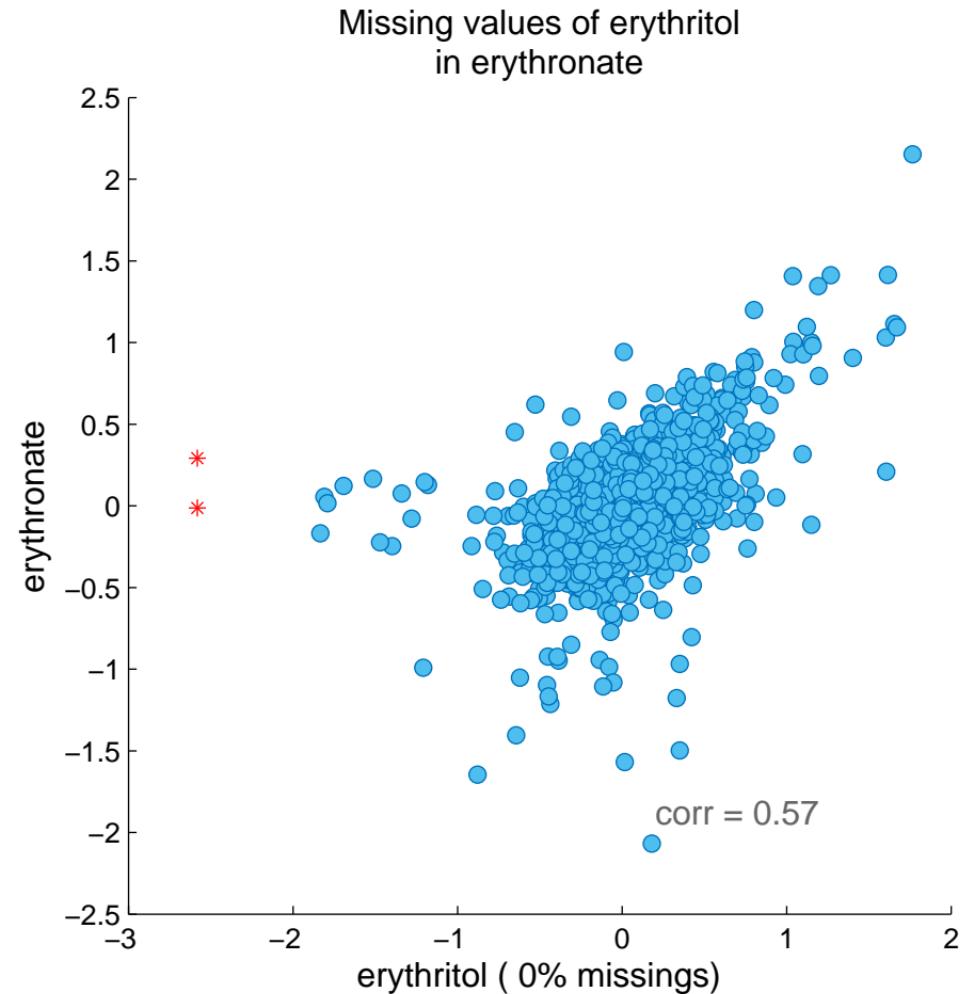


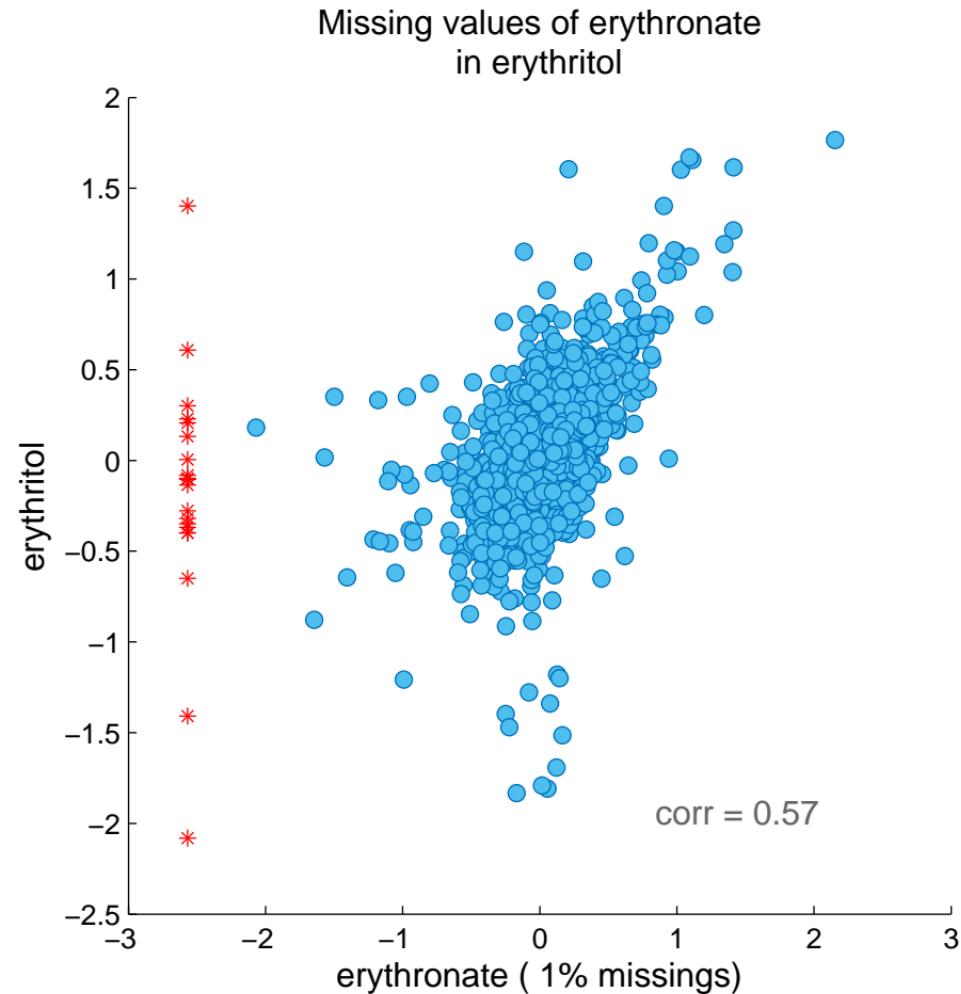




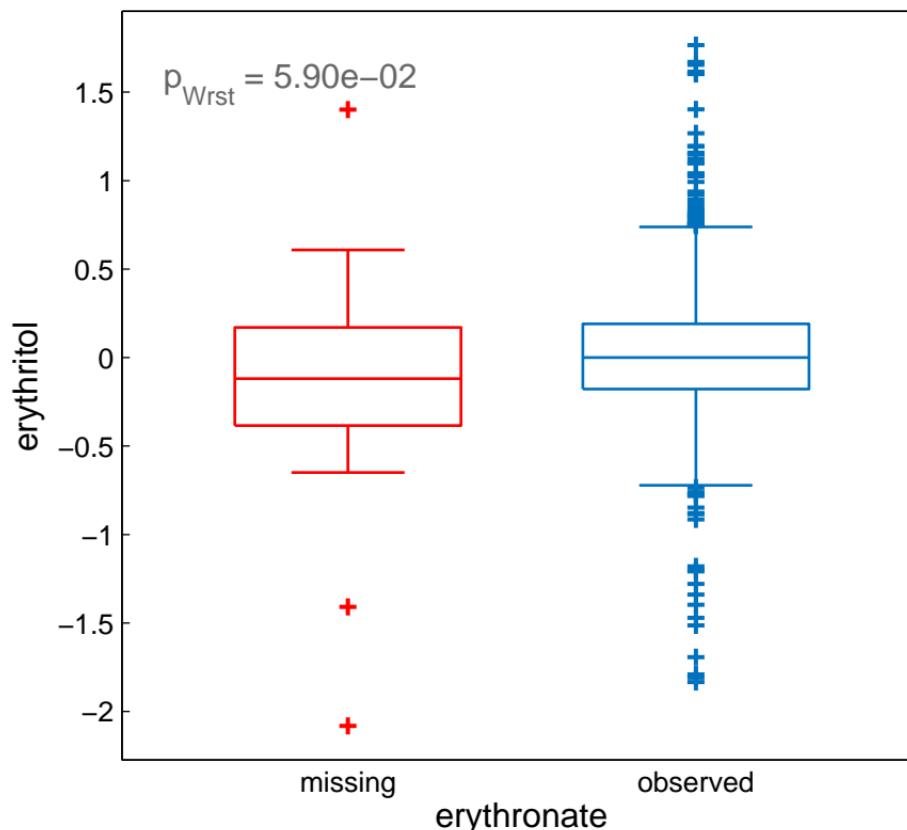


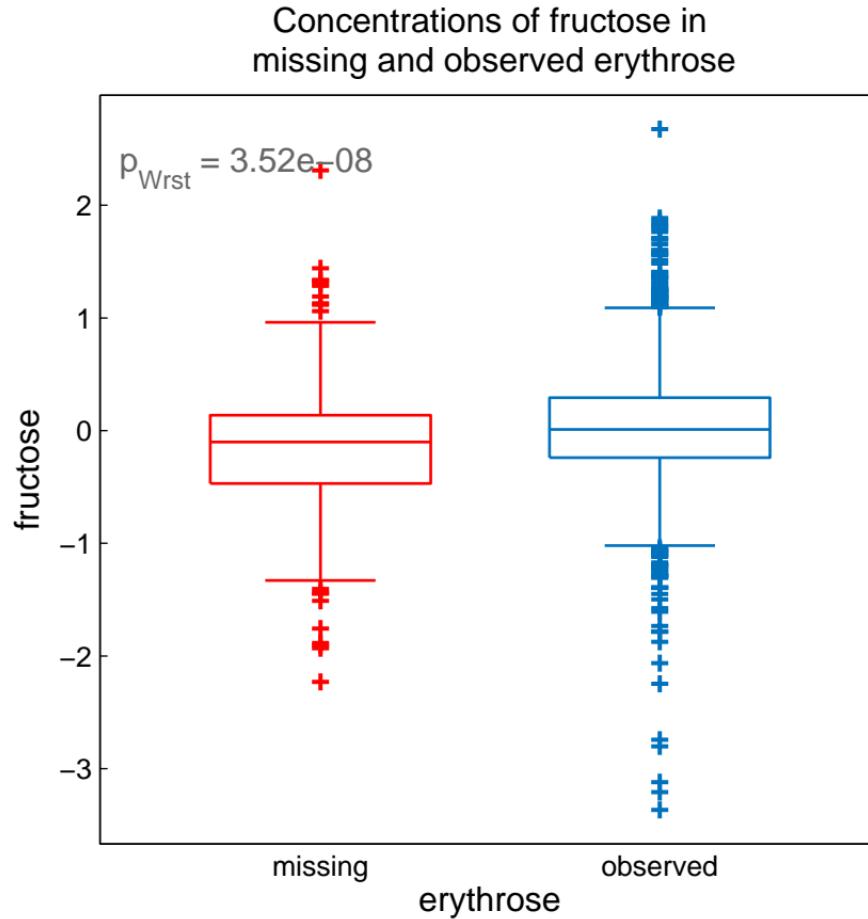
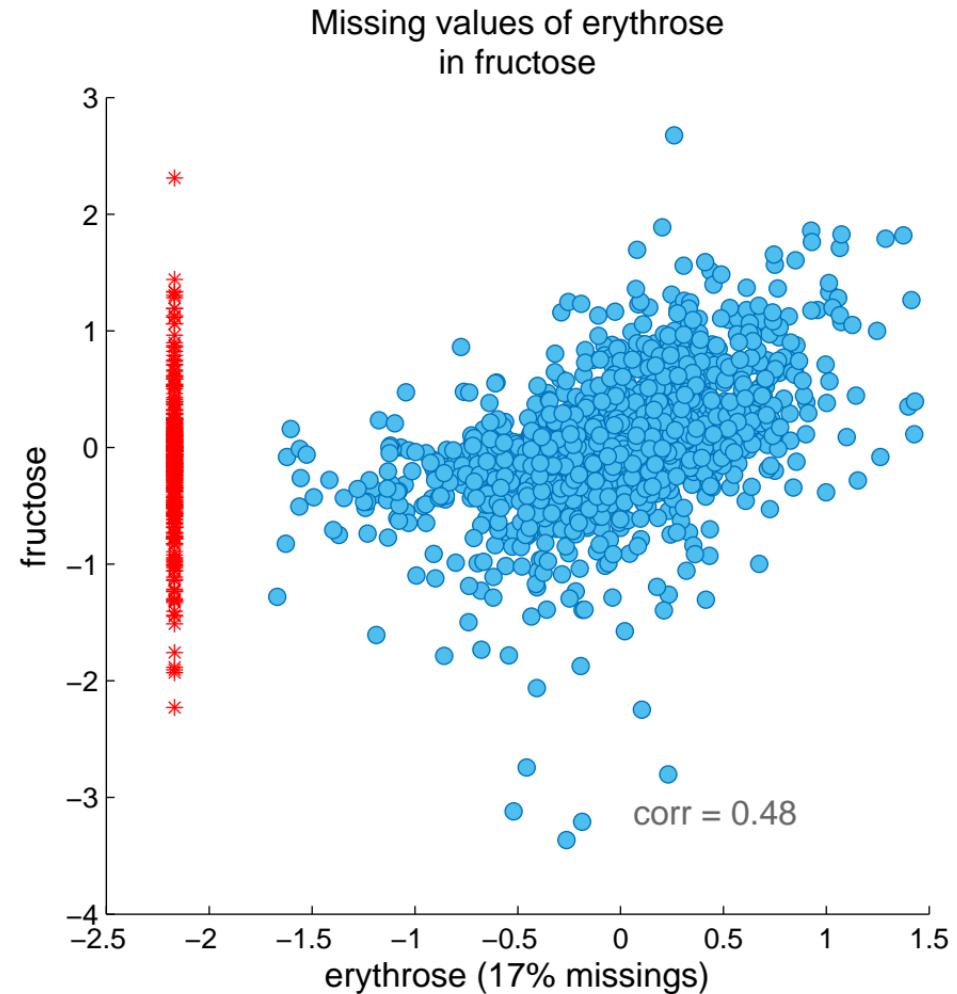


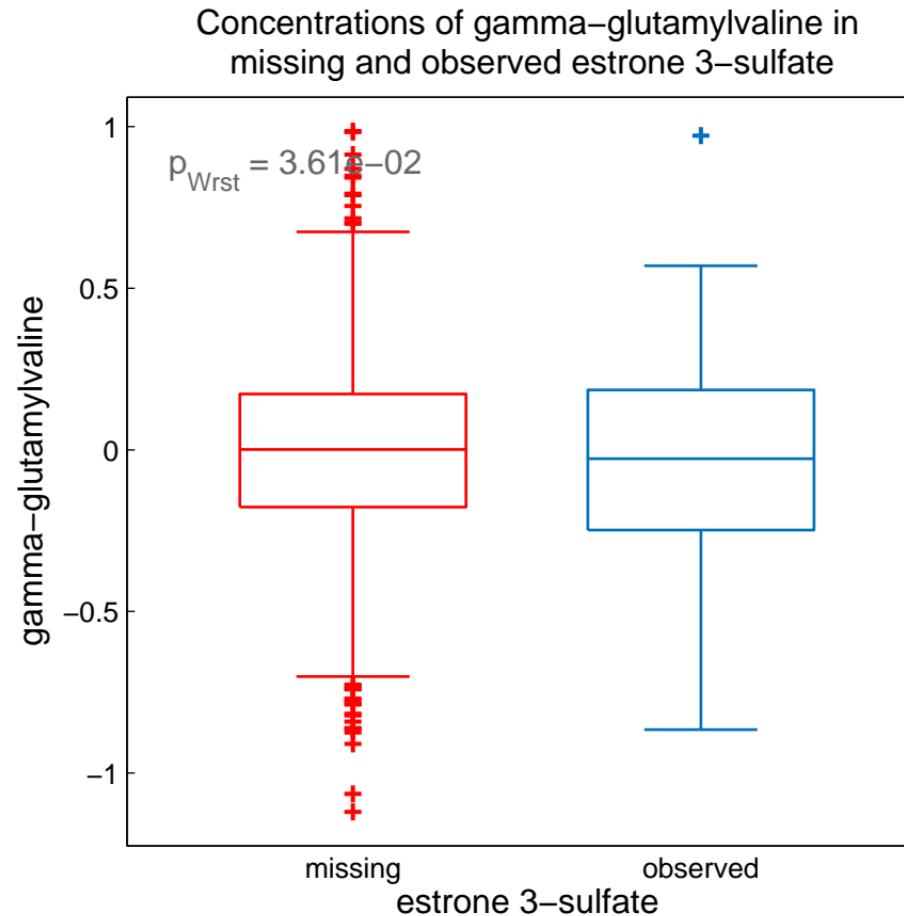
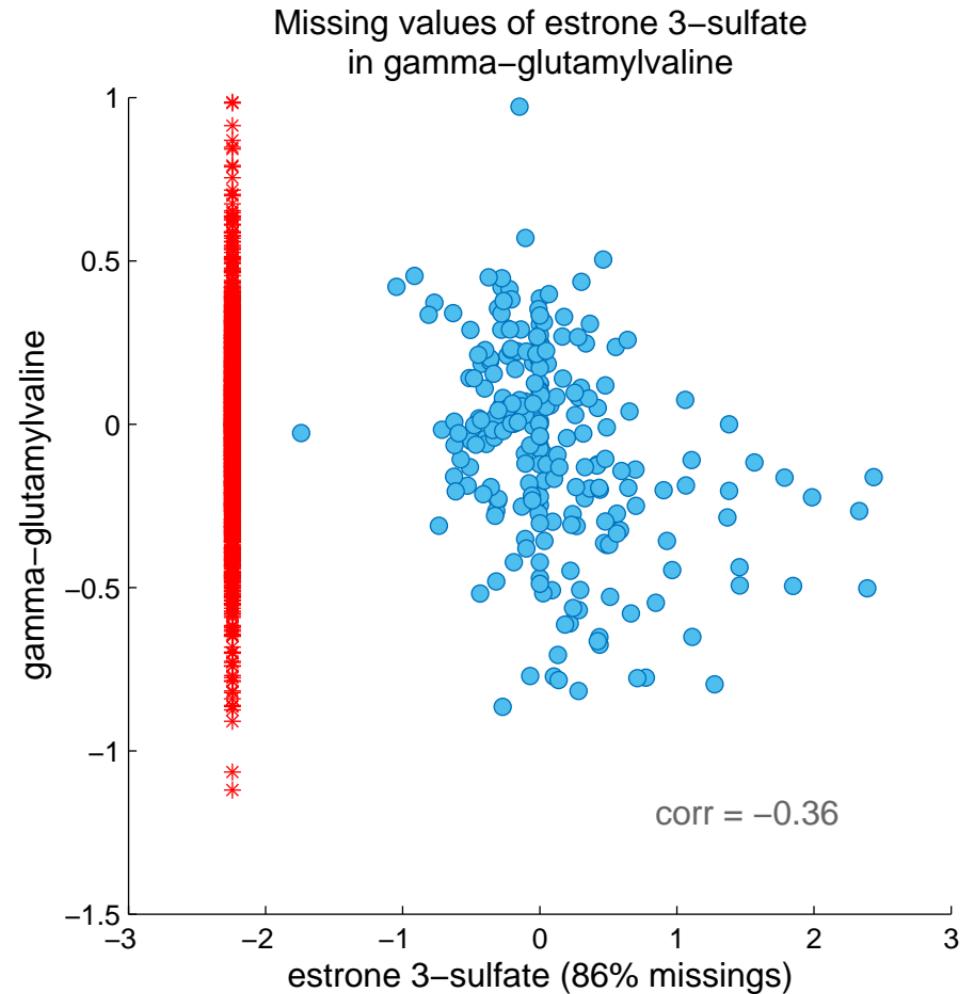


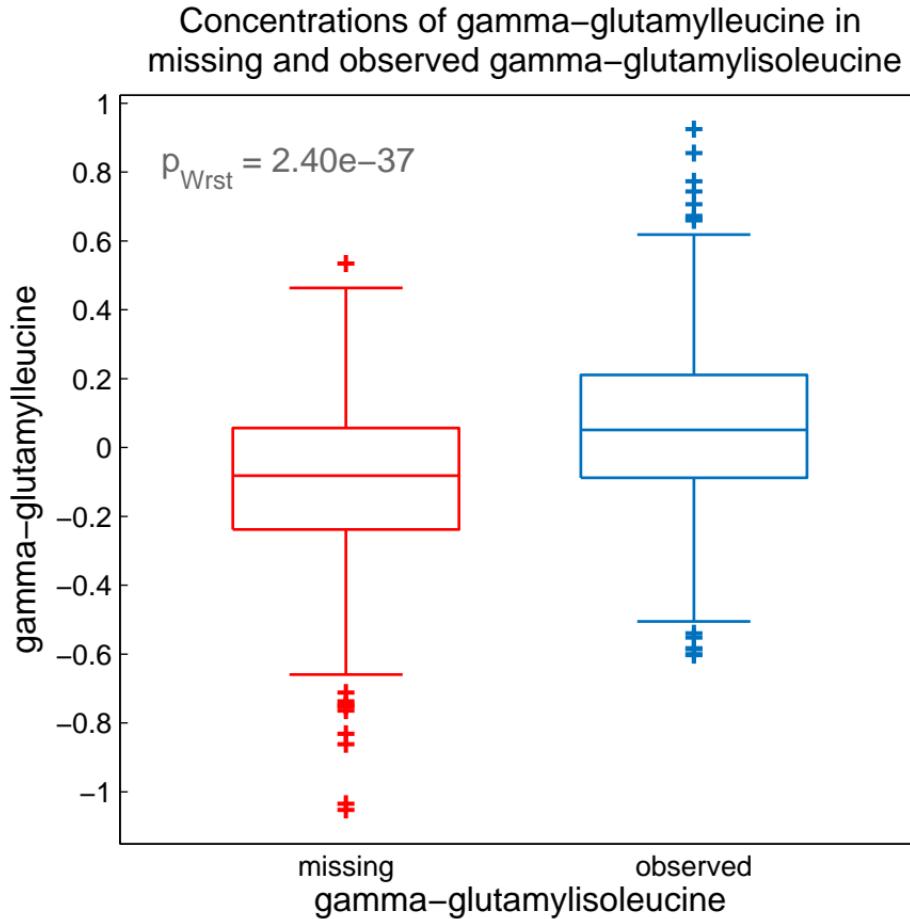
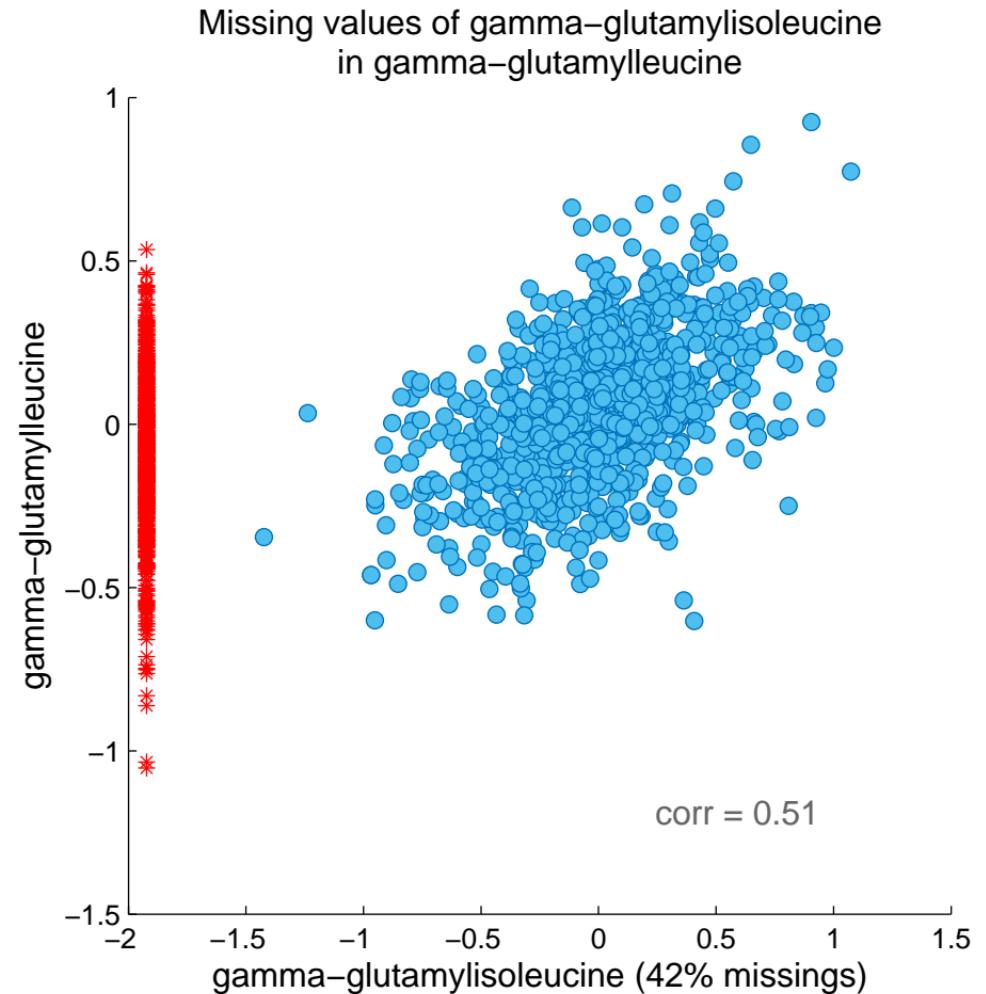


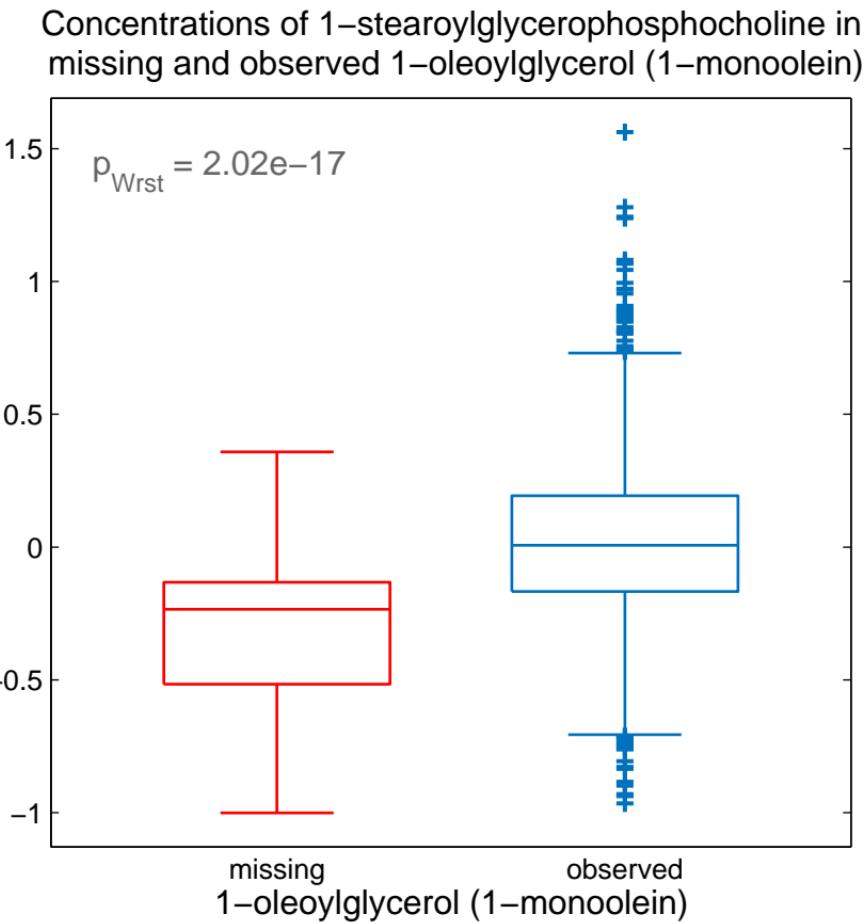
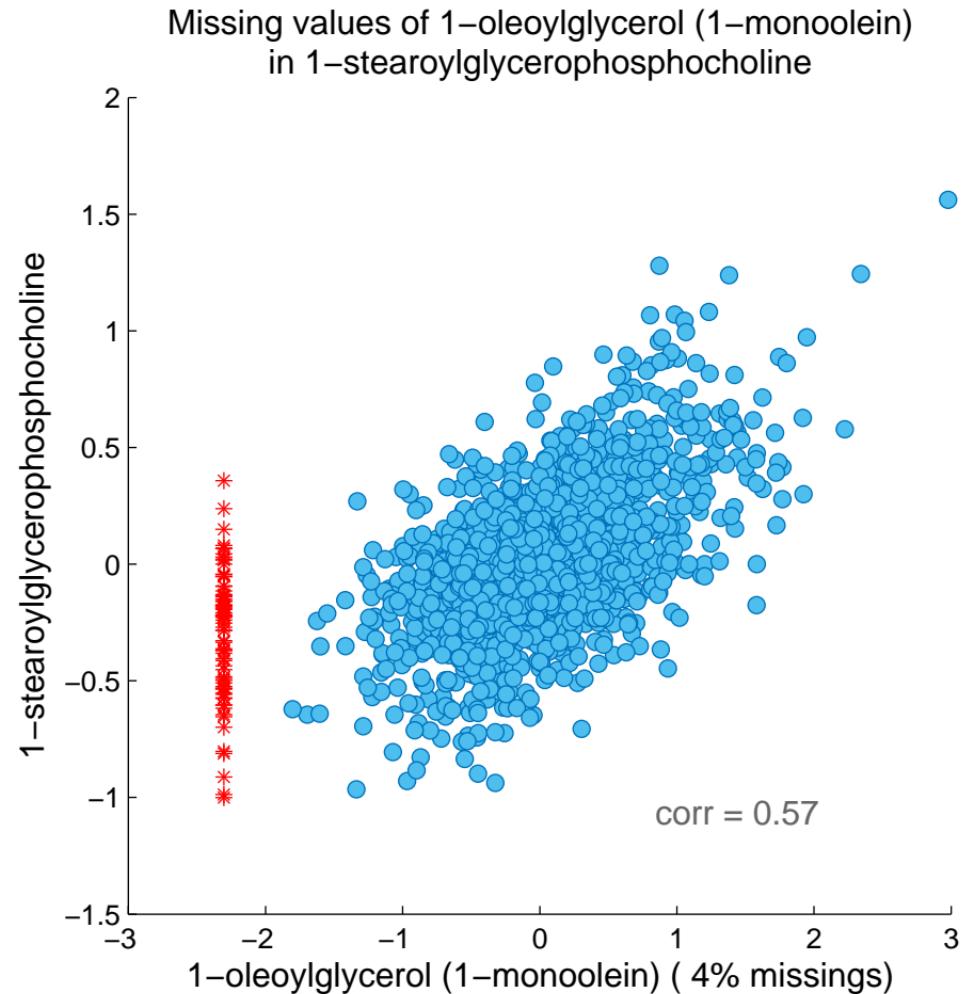
Concentrations of erythritol in
missing and observed erythronate

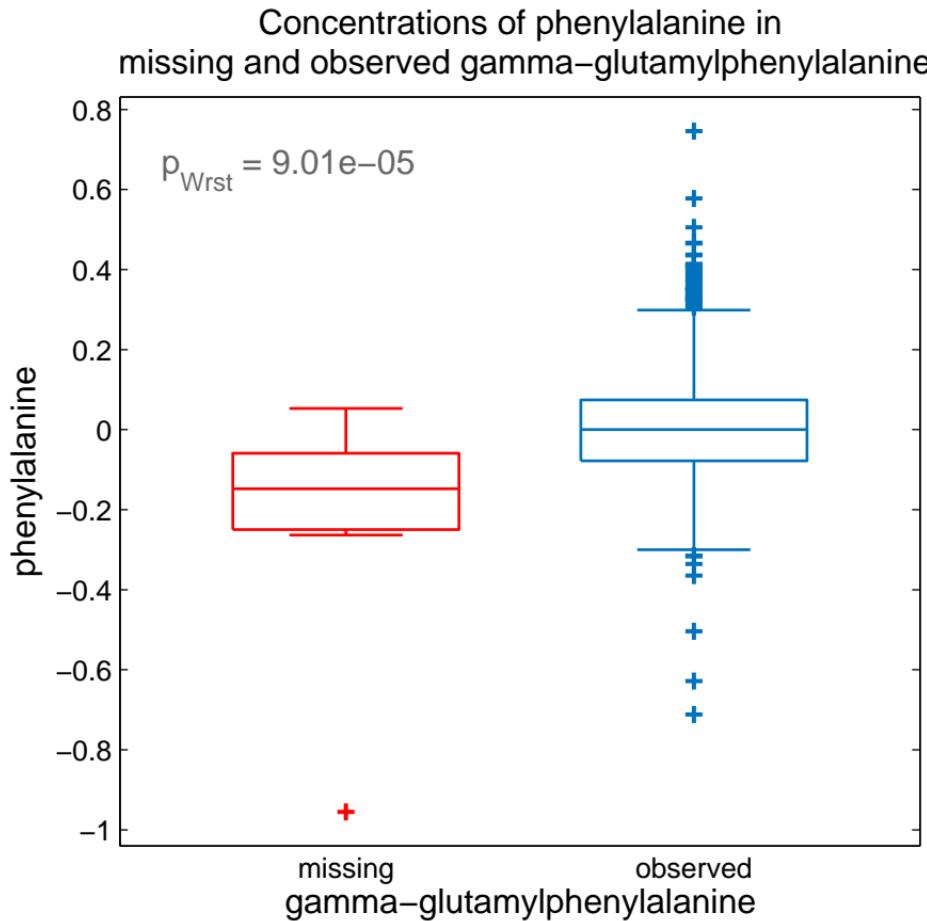
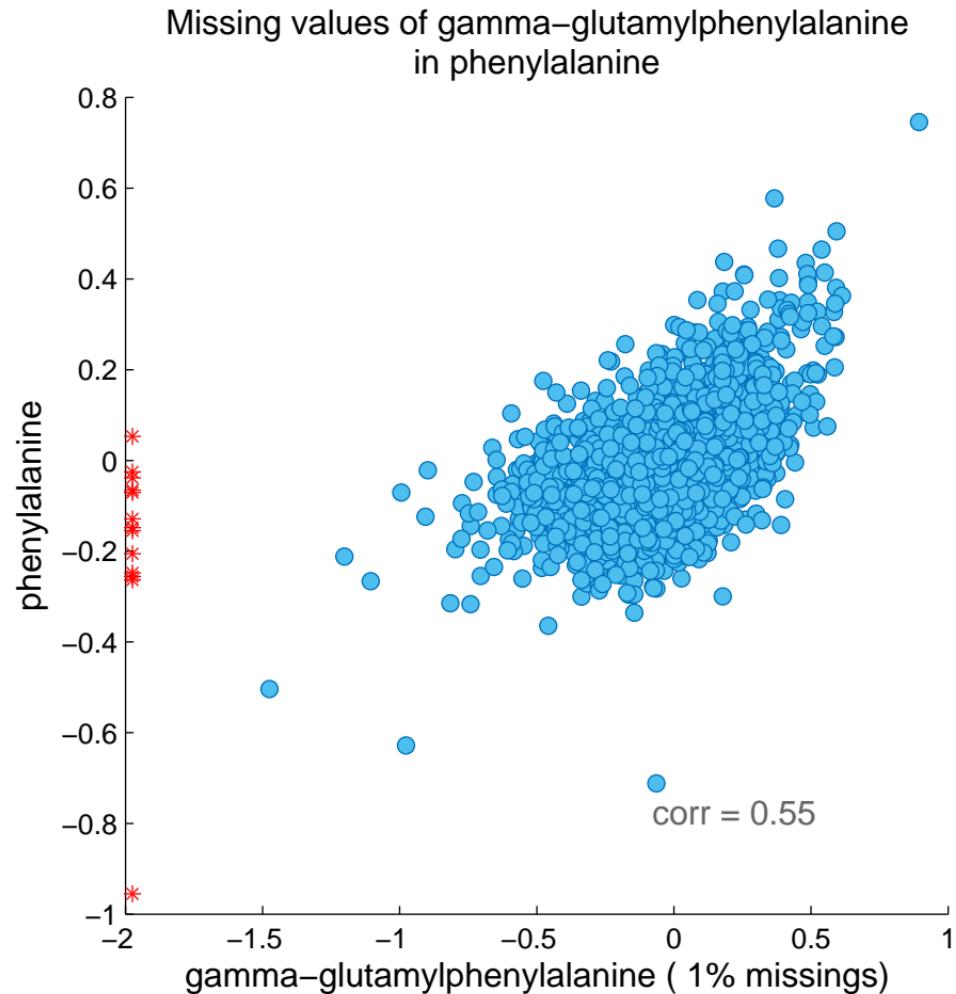


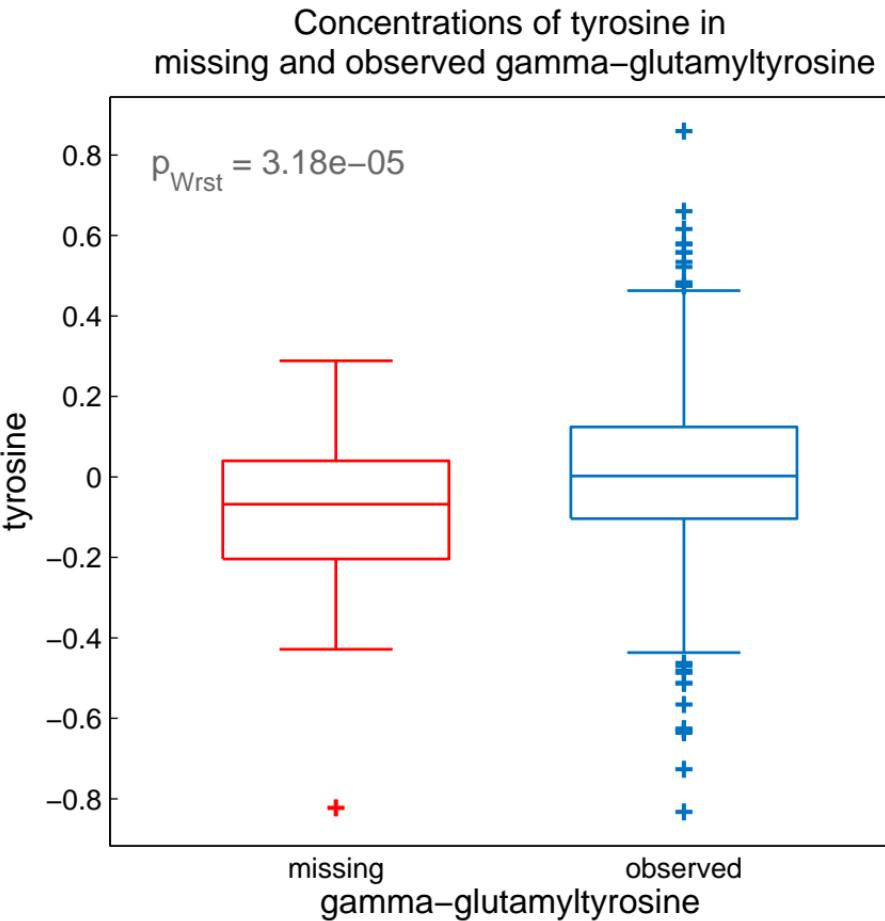
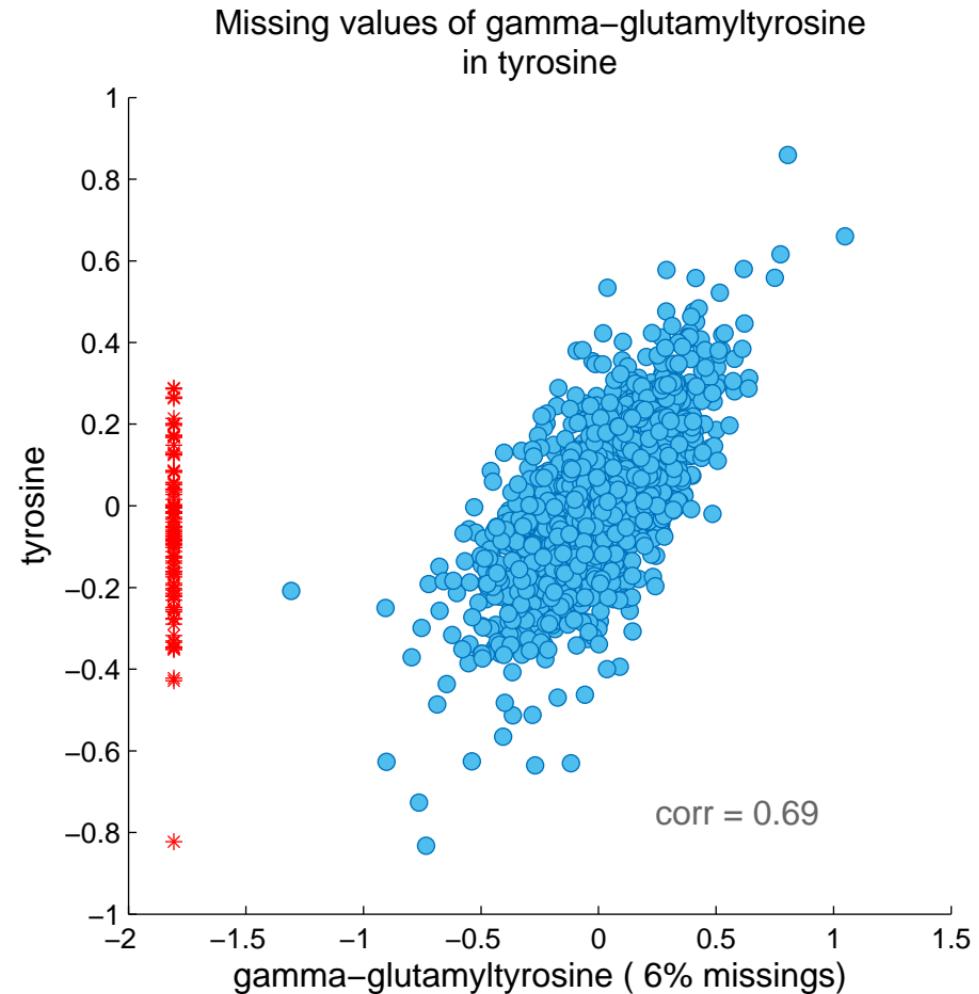




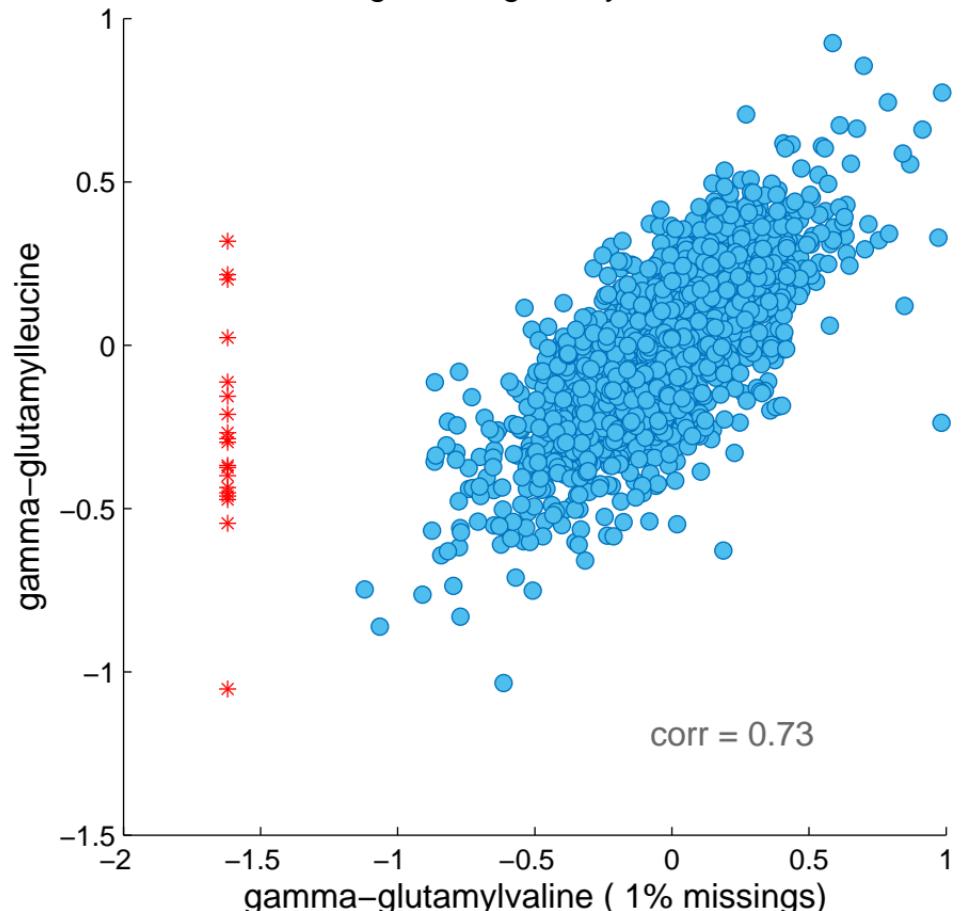




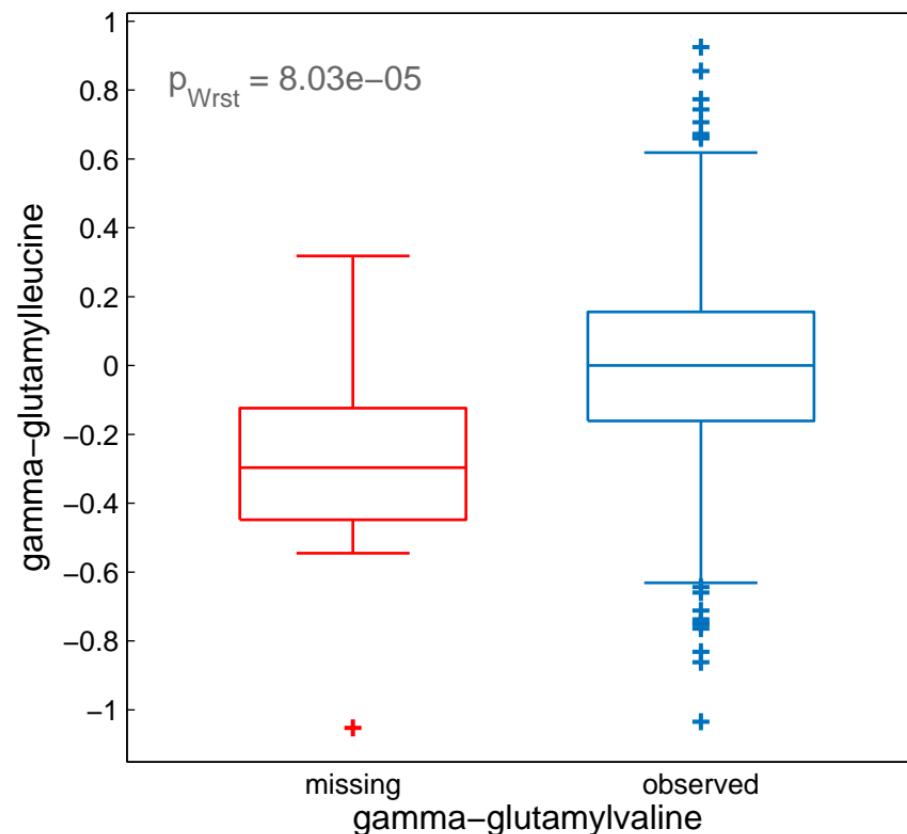


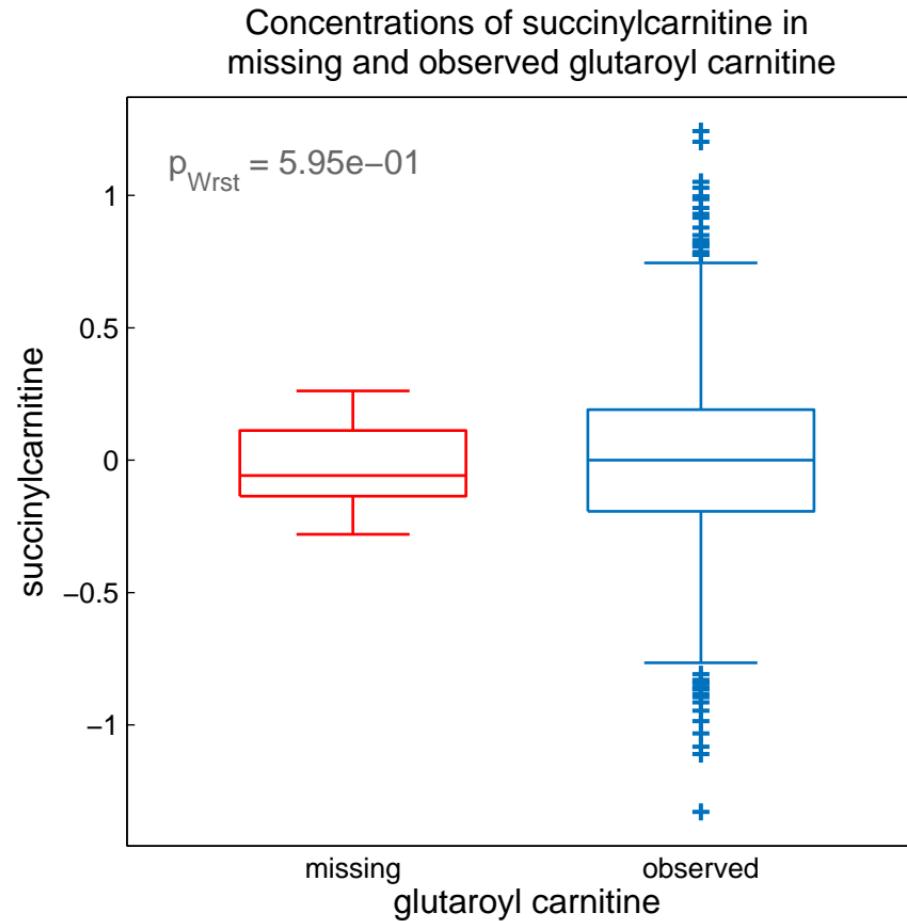
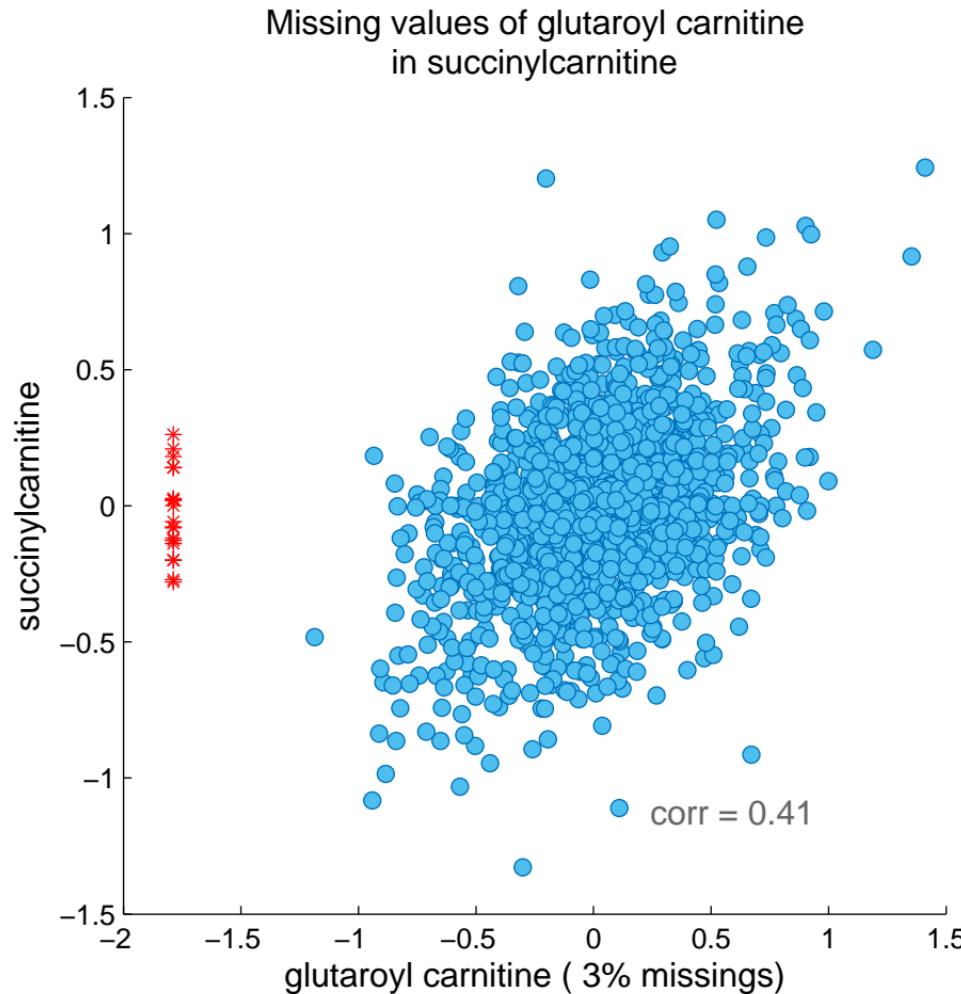


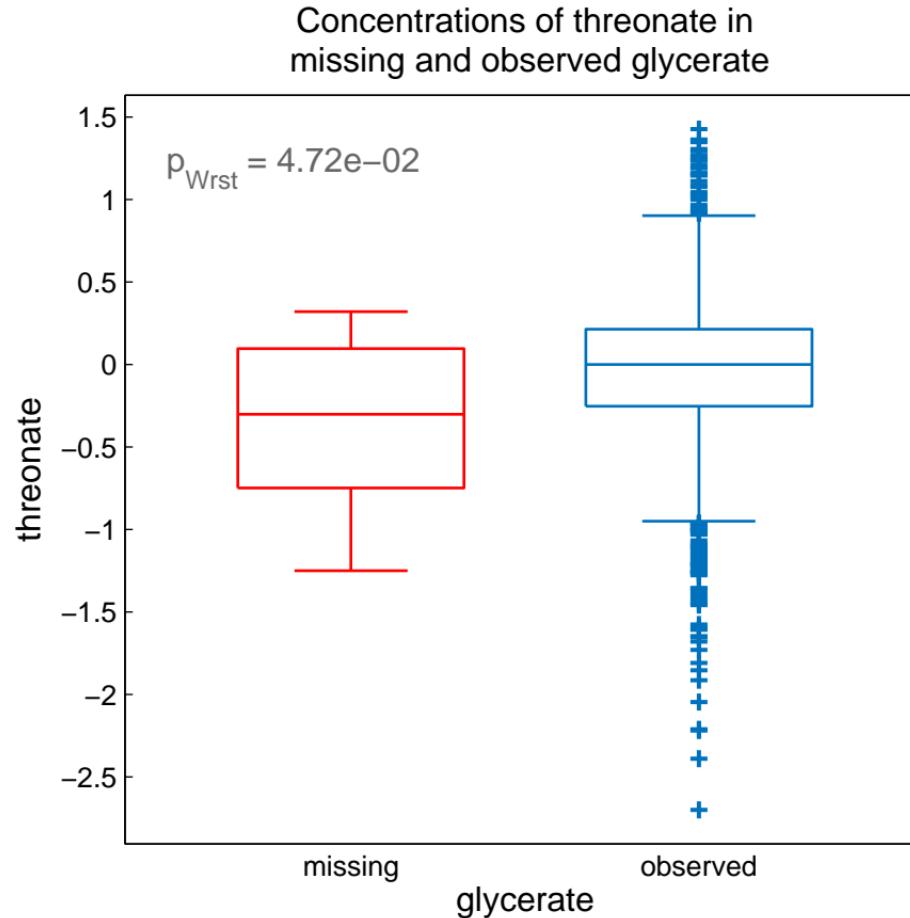
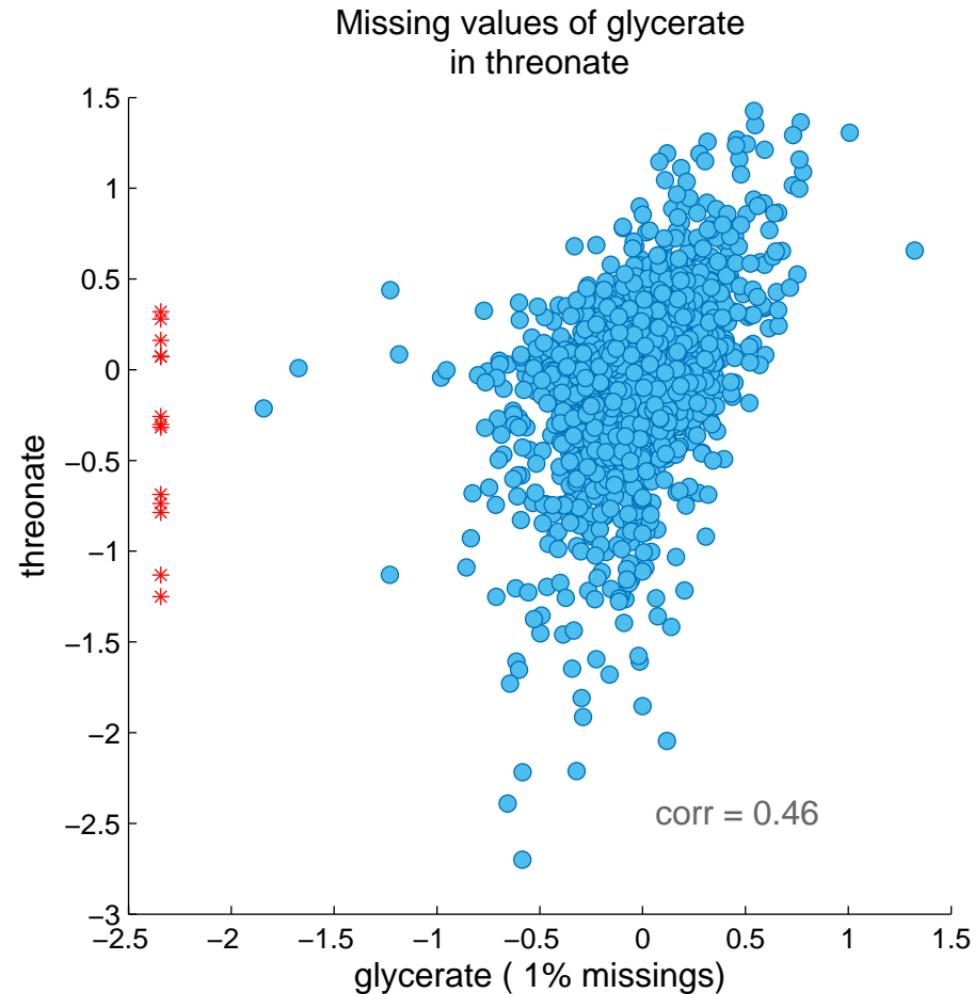
Missing values of gamma-glutamylvaline
in gamma-glutamylleucine

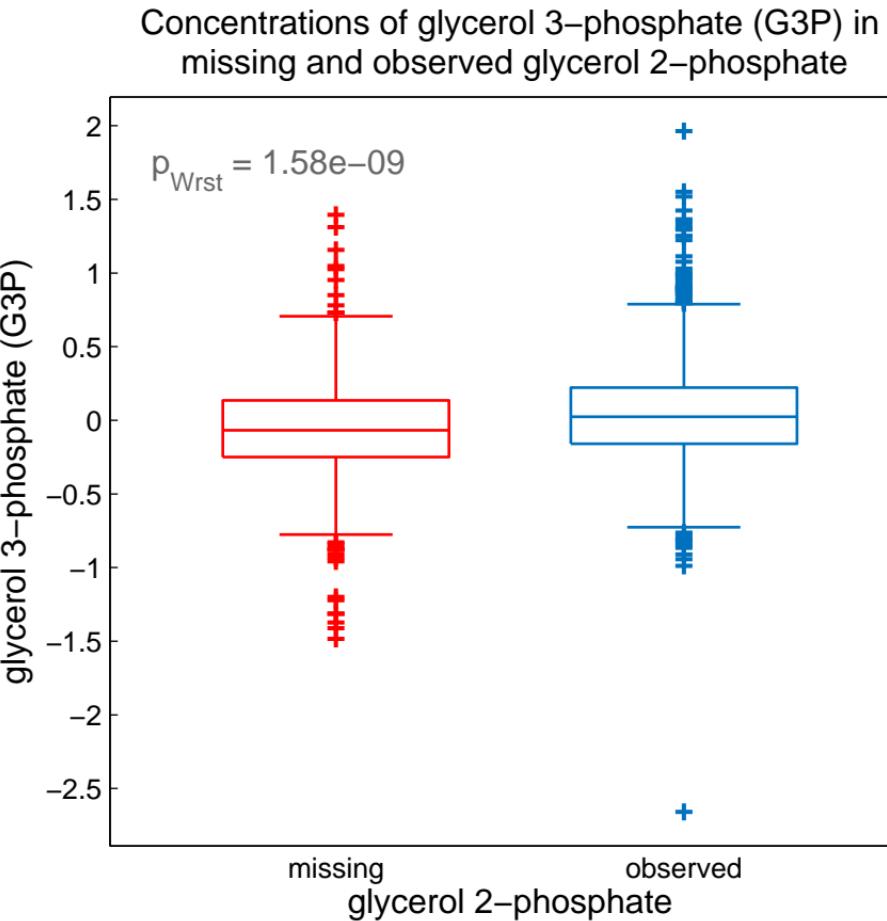
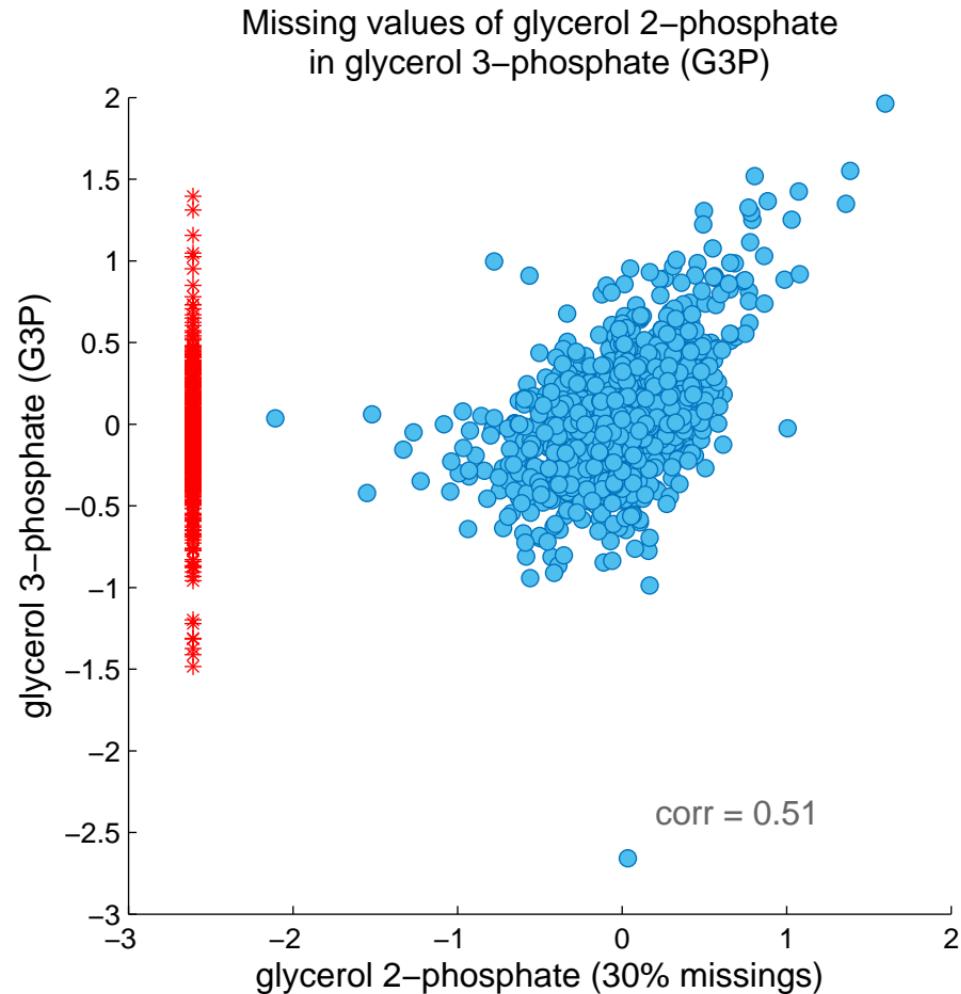


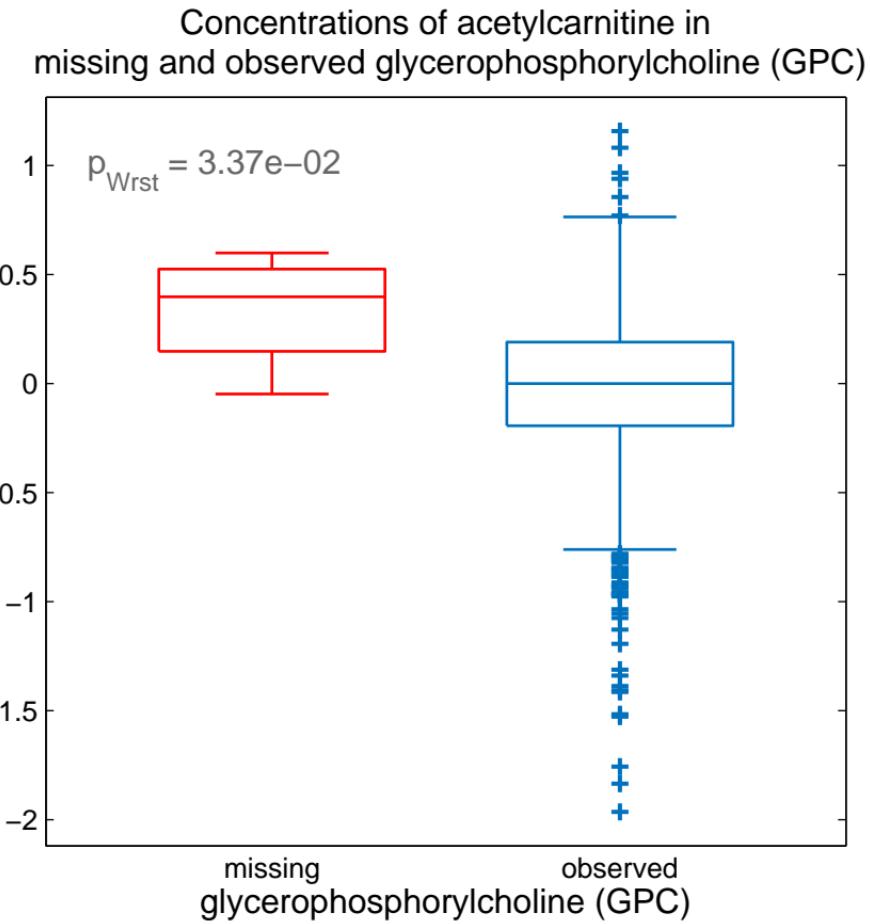
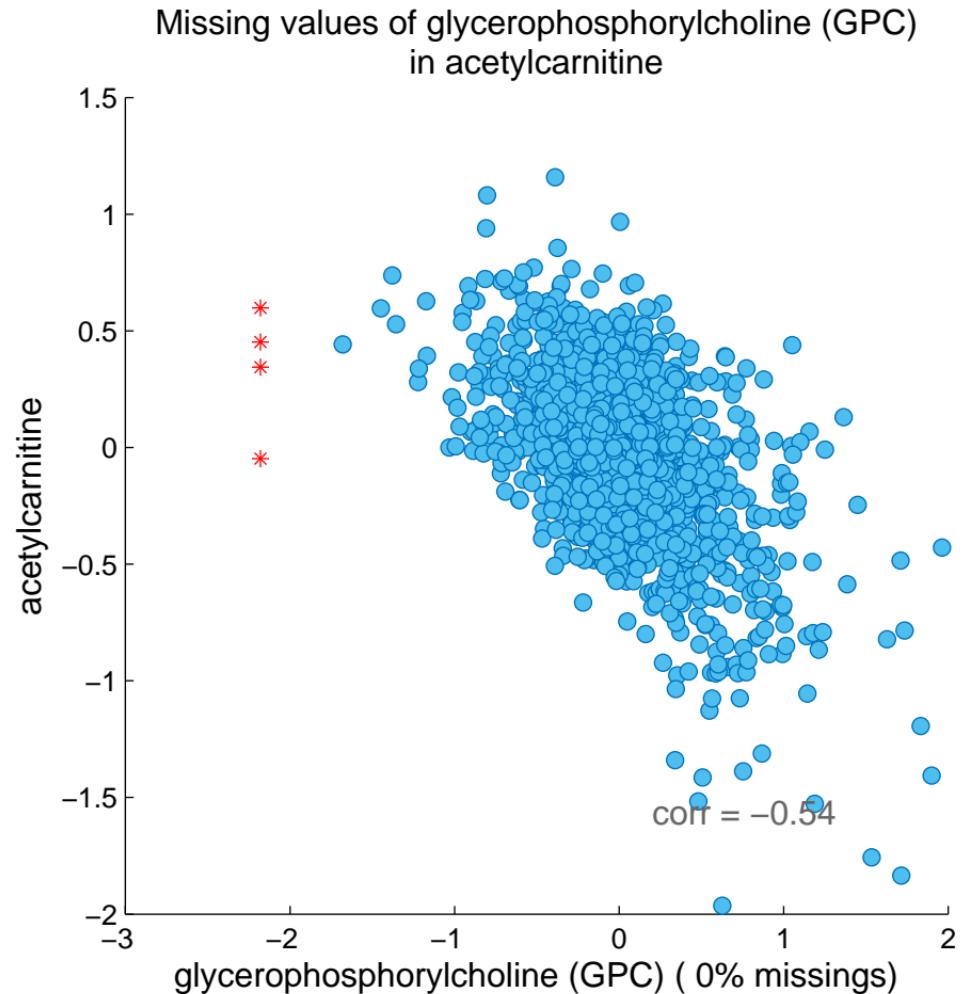
Concentrations of gamma-glutamylleucine in
missing and observed gamma-glutamylvaline



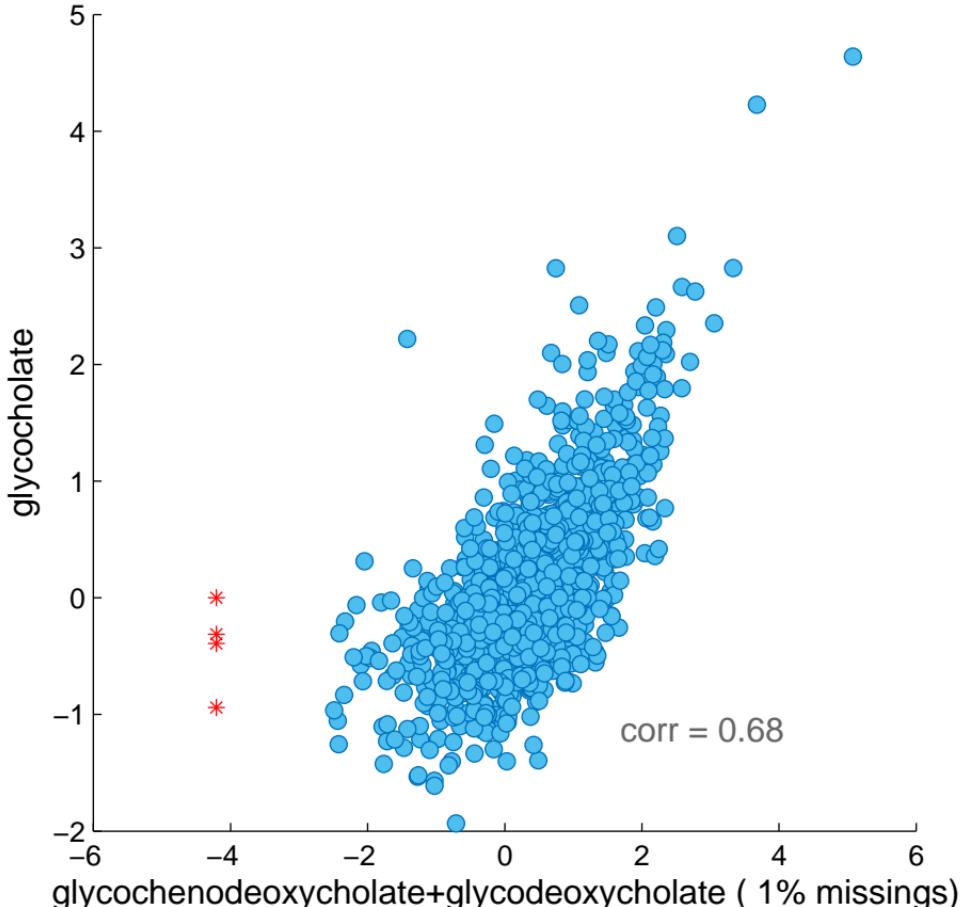




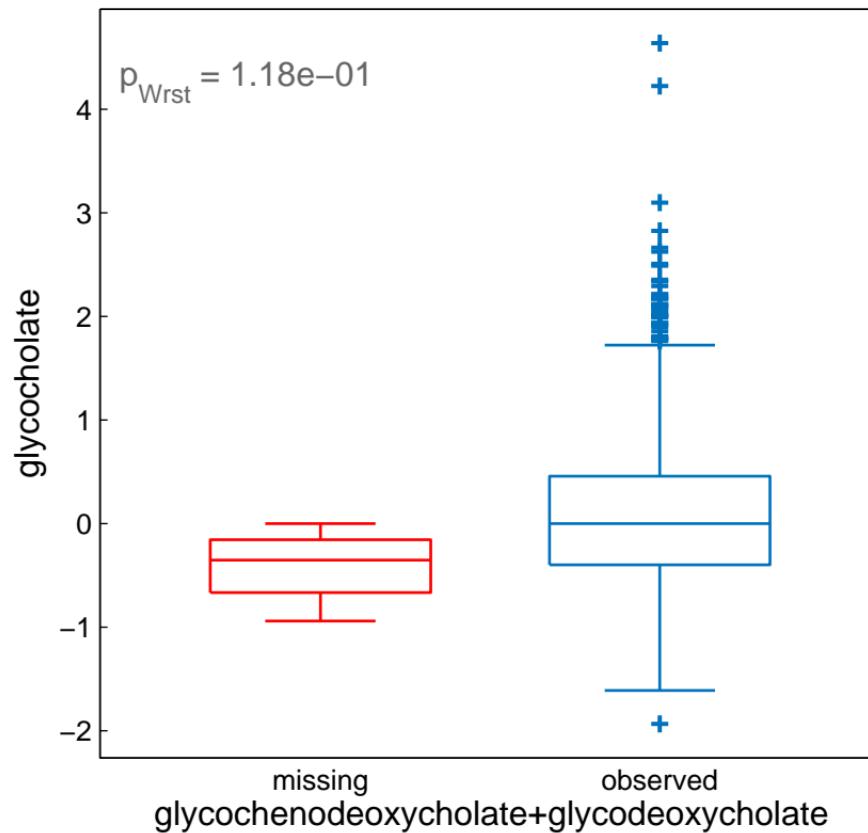


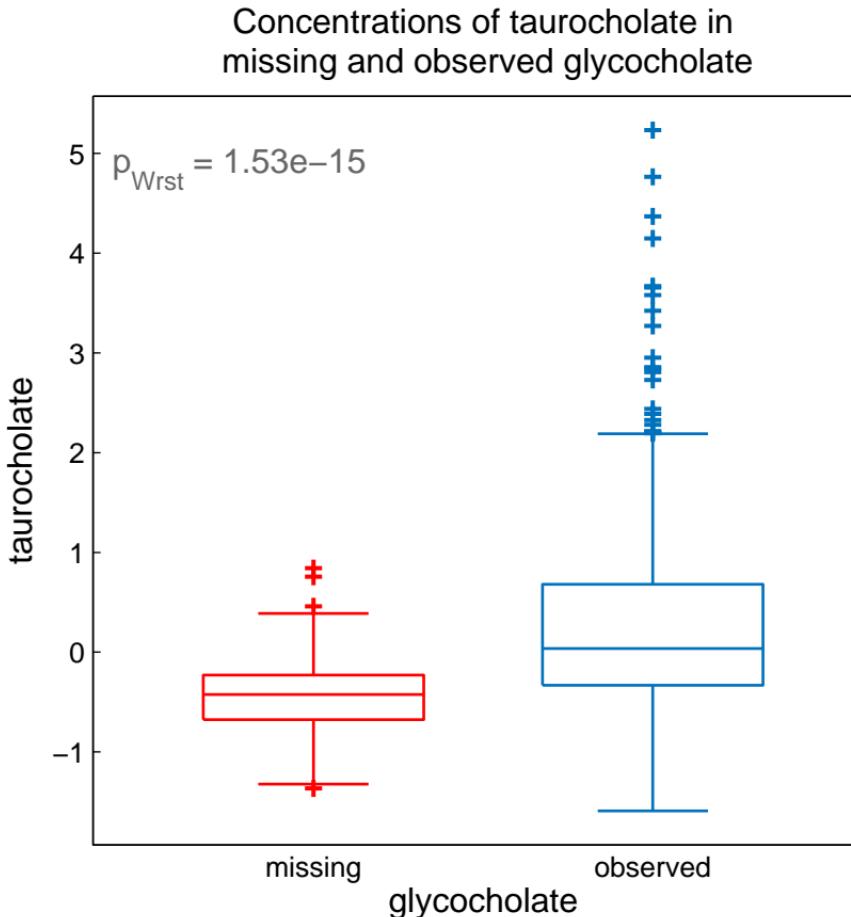
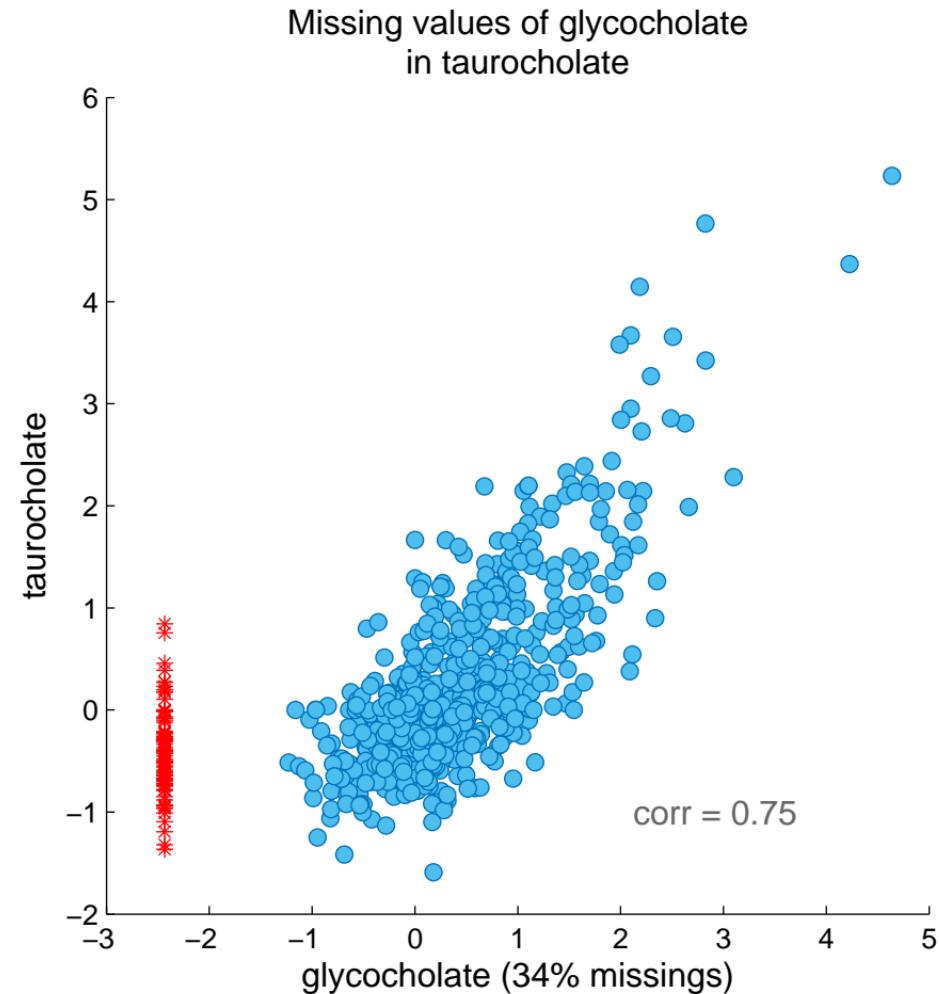


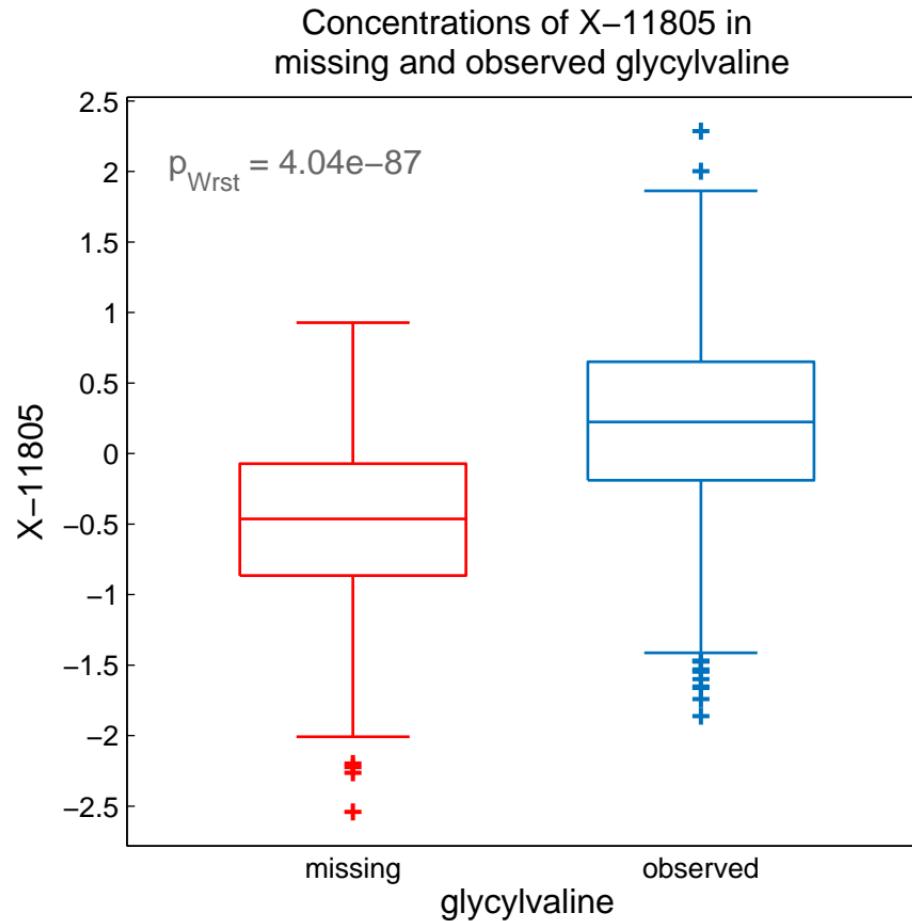
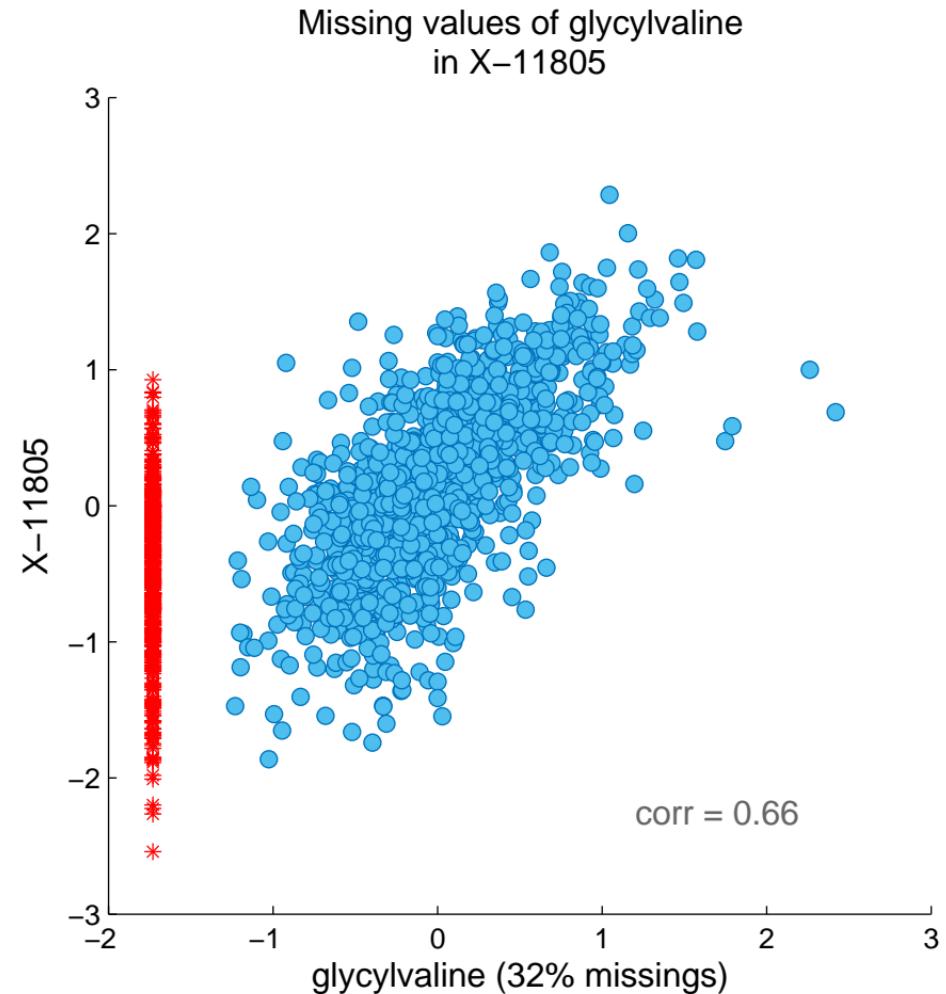
Missing values of glycochenodeoxycholate+glycodeoxycholate
in glycocholate

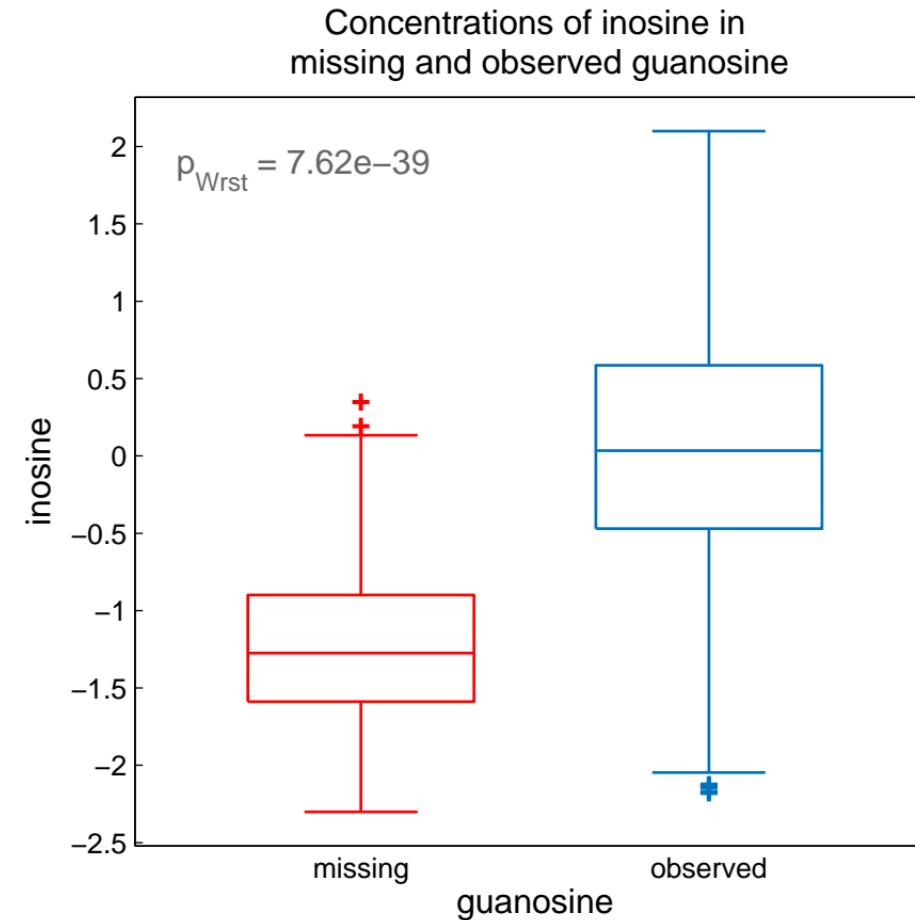
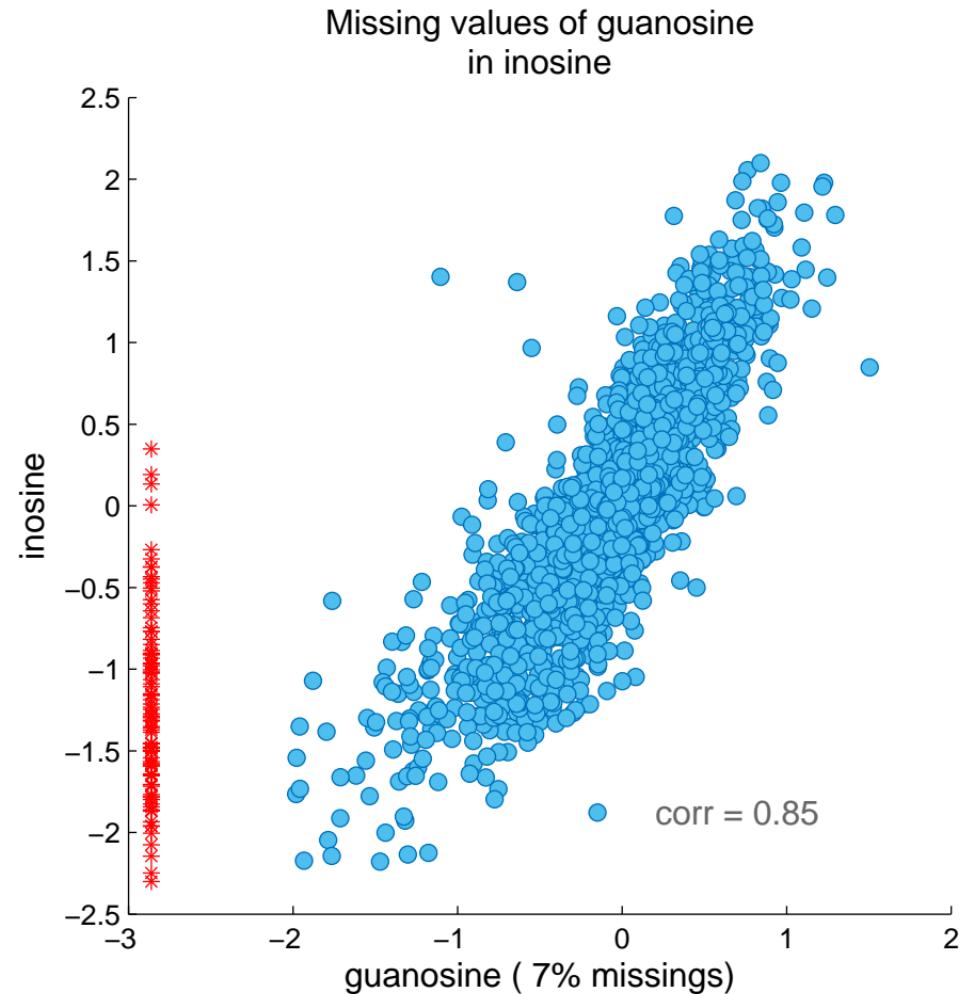


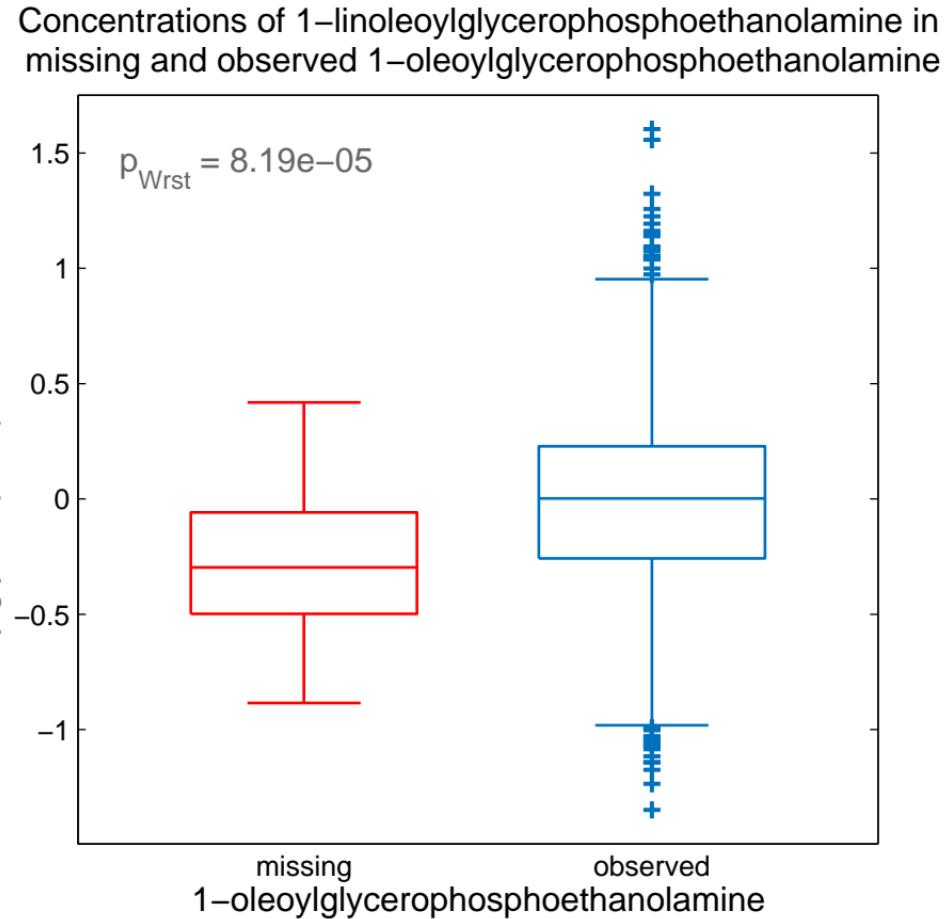
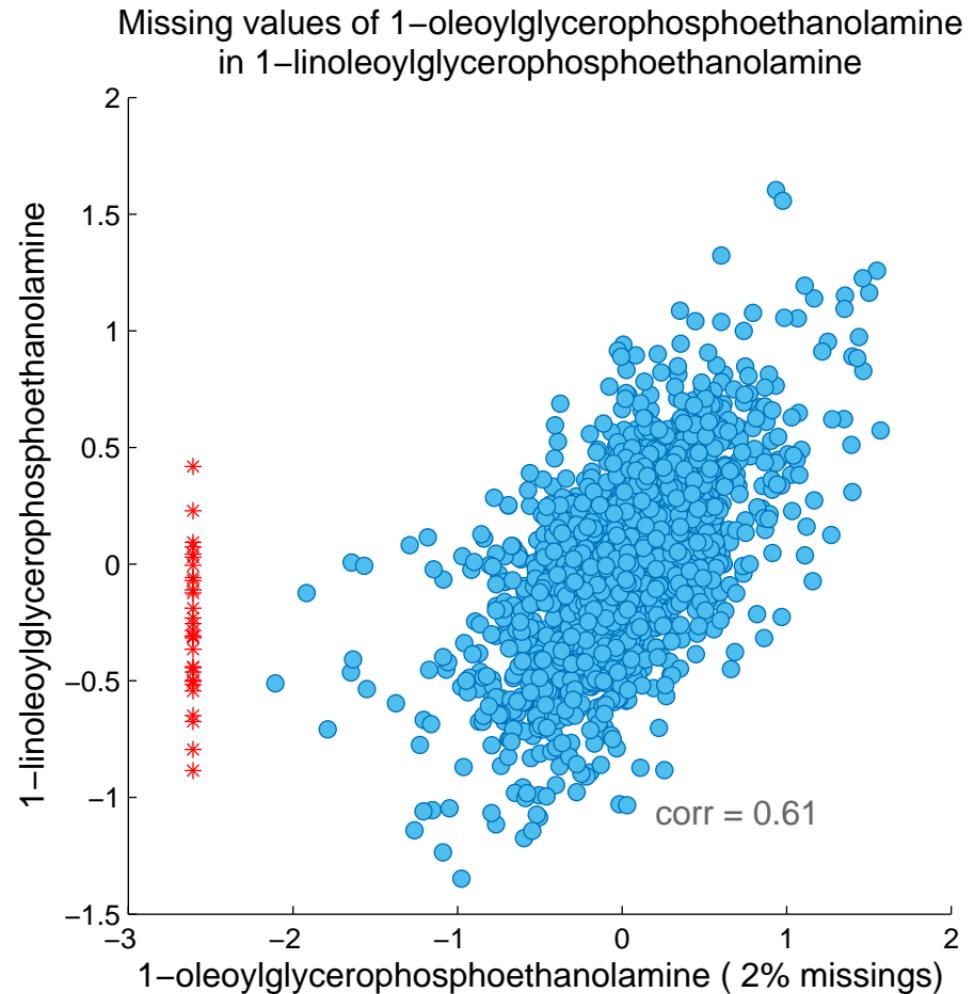
Concentrations of glycocholate in
missing and observed glycochenodeoxycholate+glycodeoxycholate



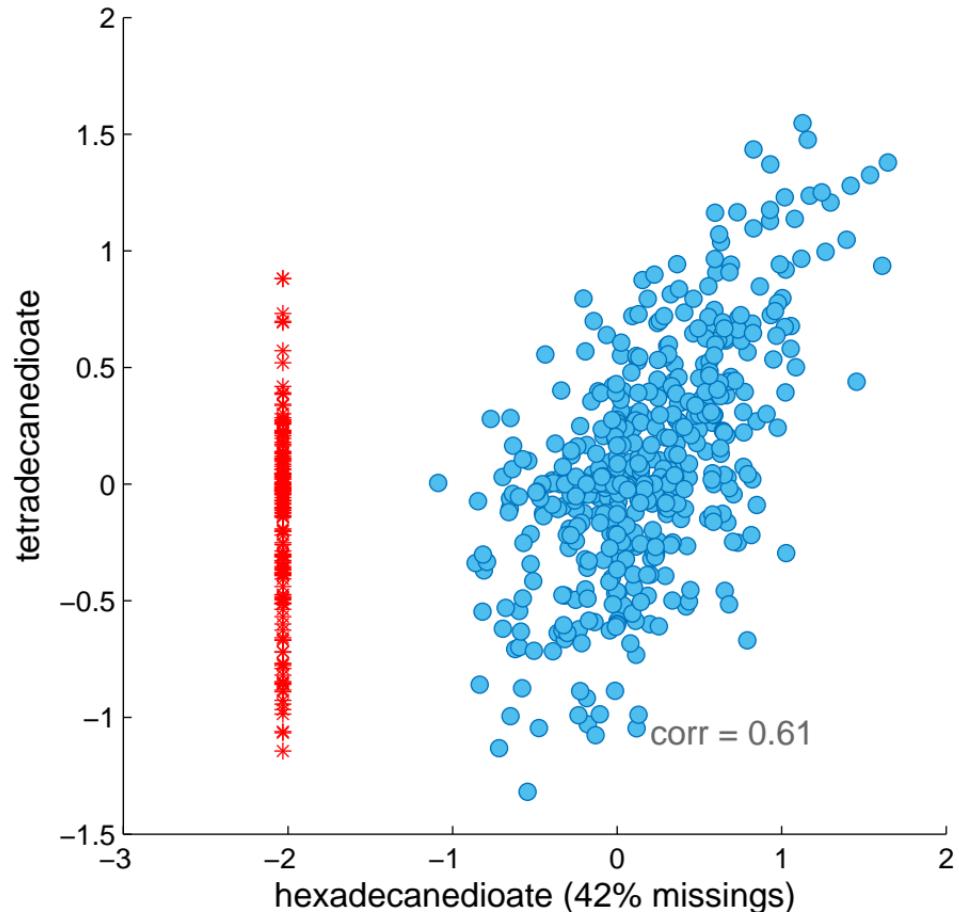




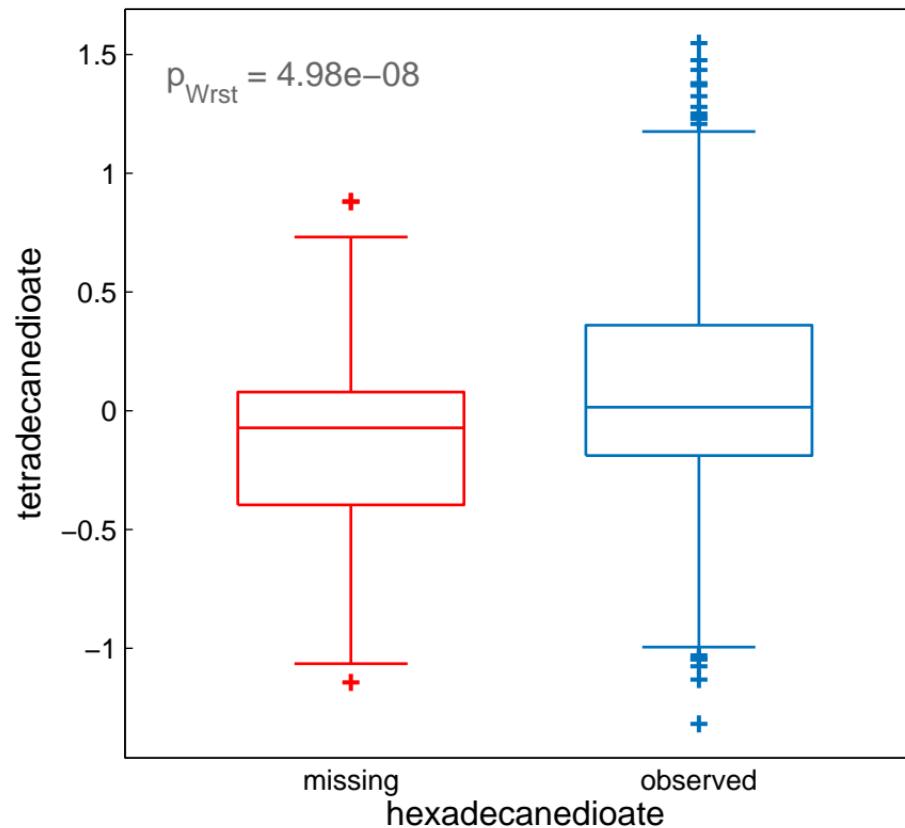


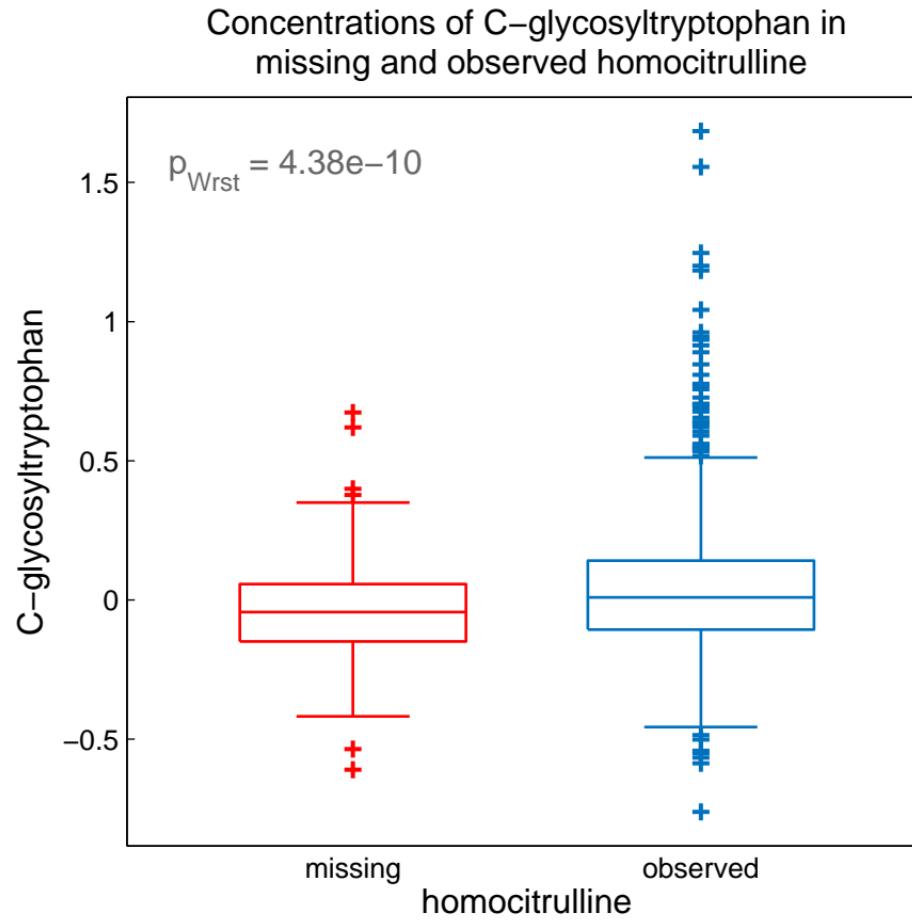
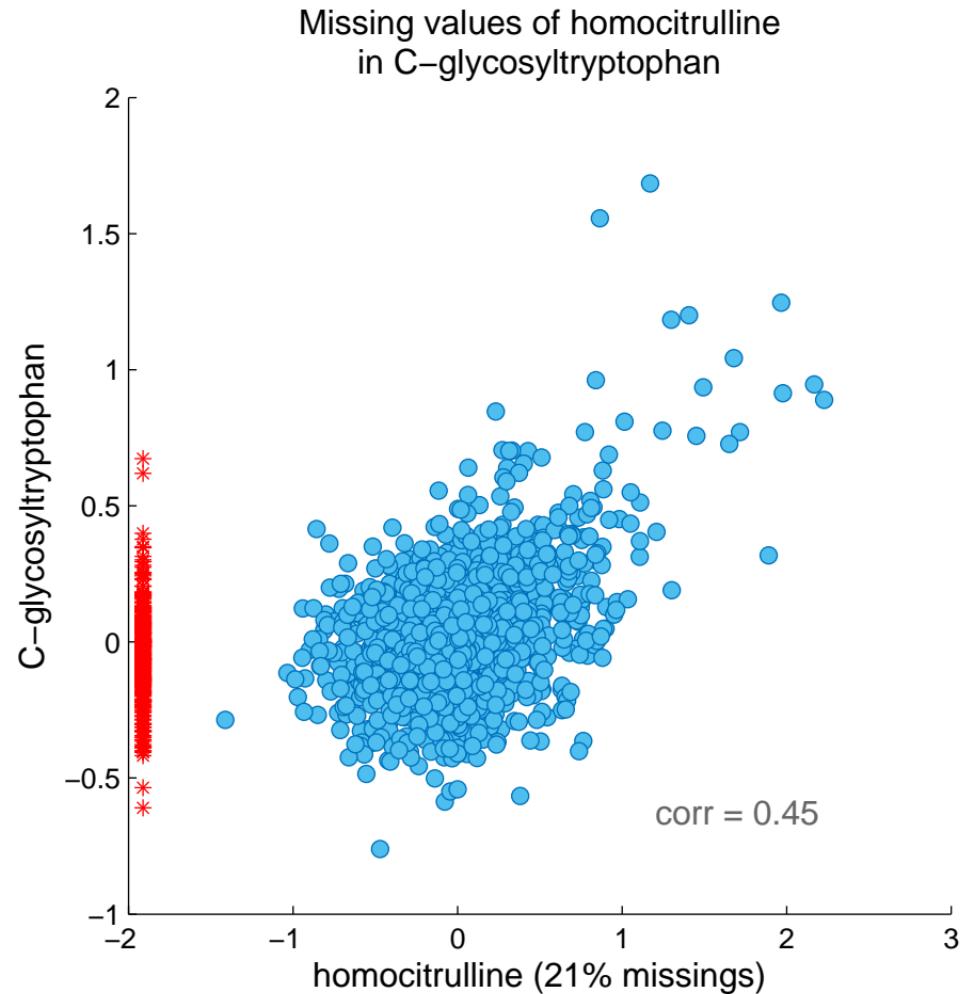


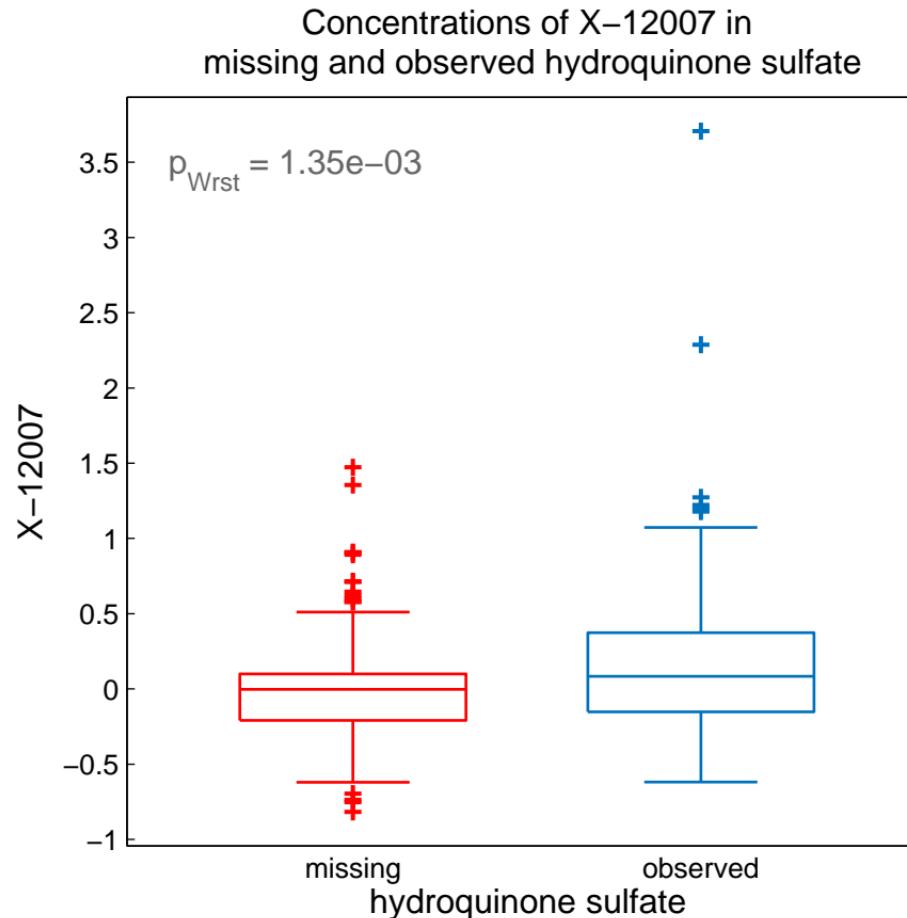
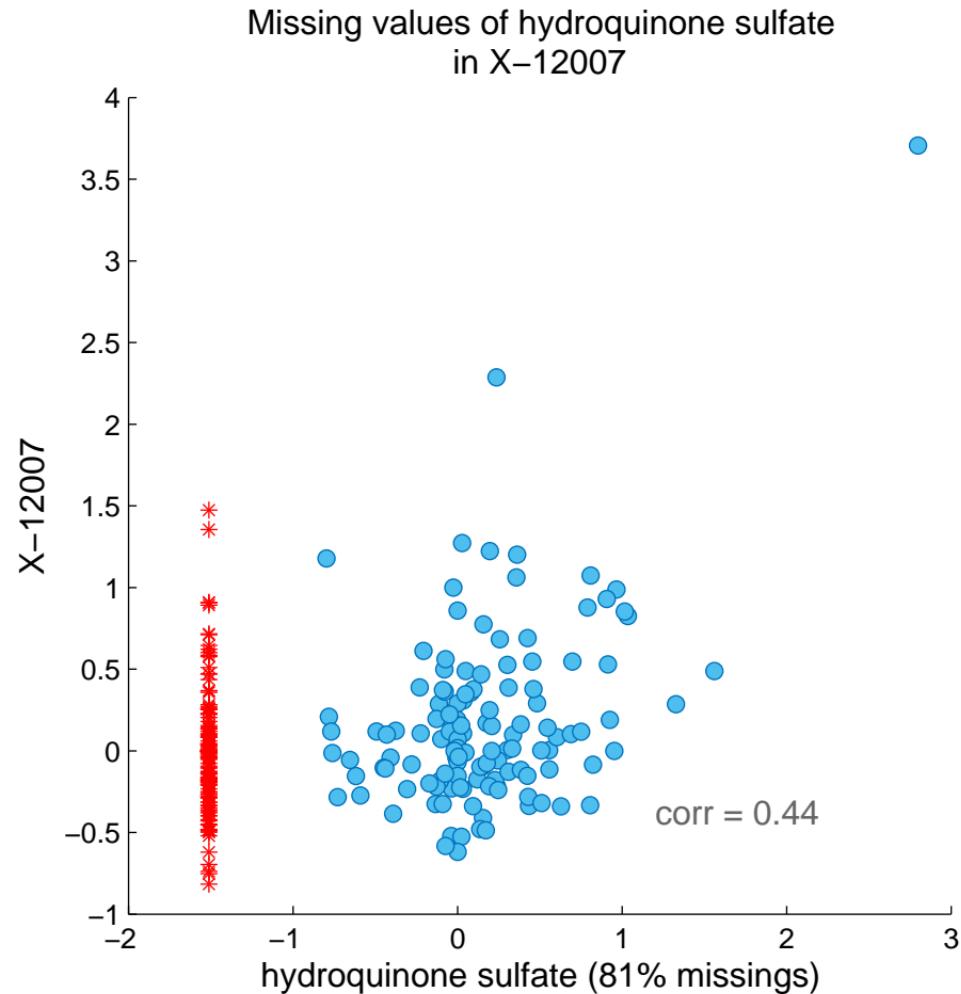
Missing values of hexadecanedioate
in tetradecanedioate

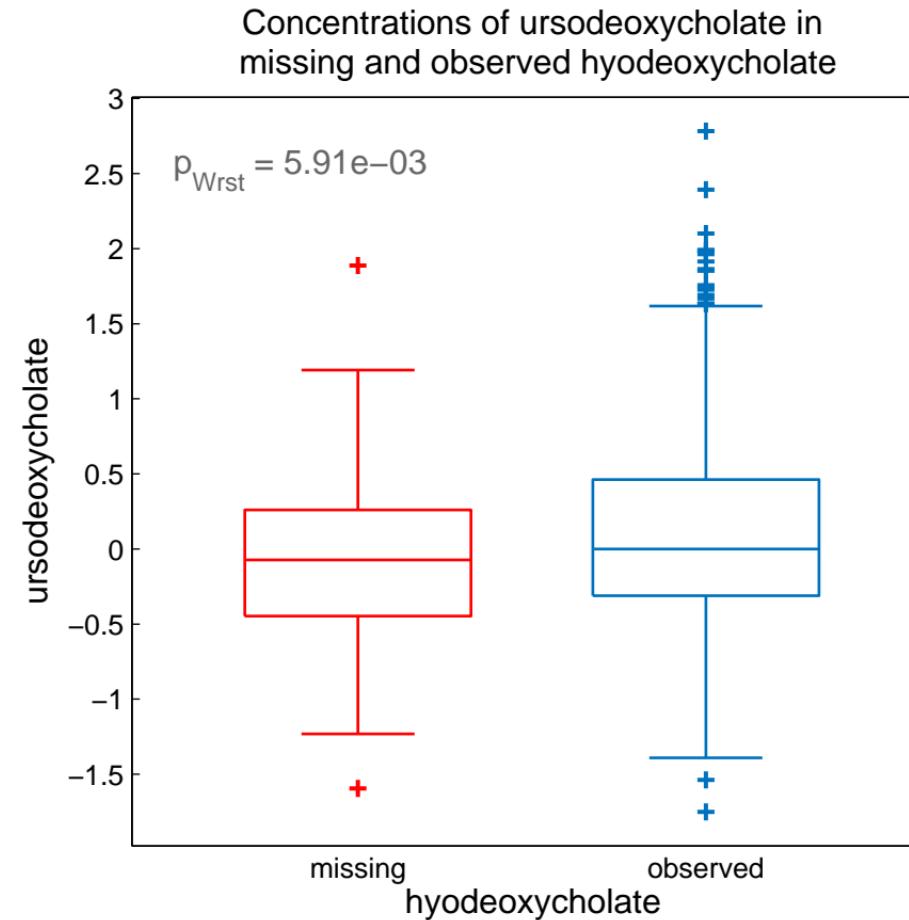
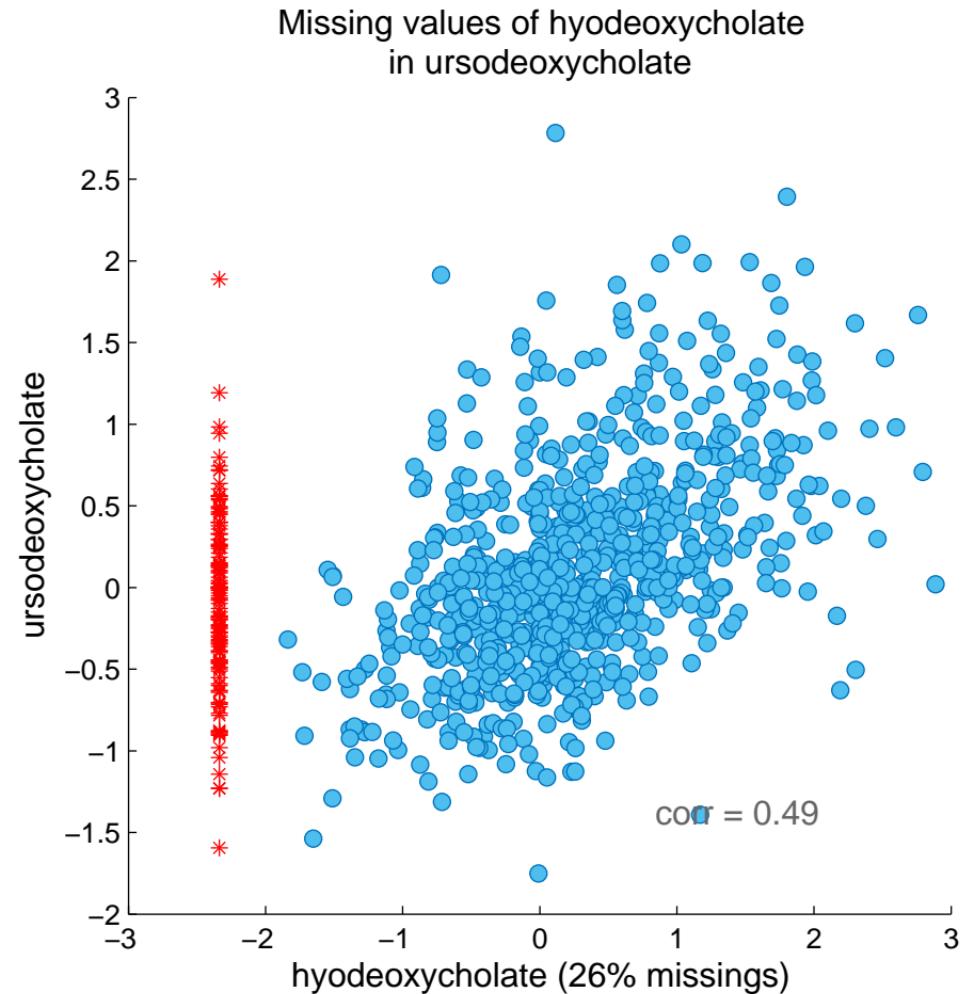


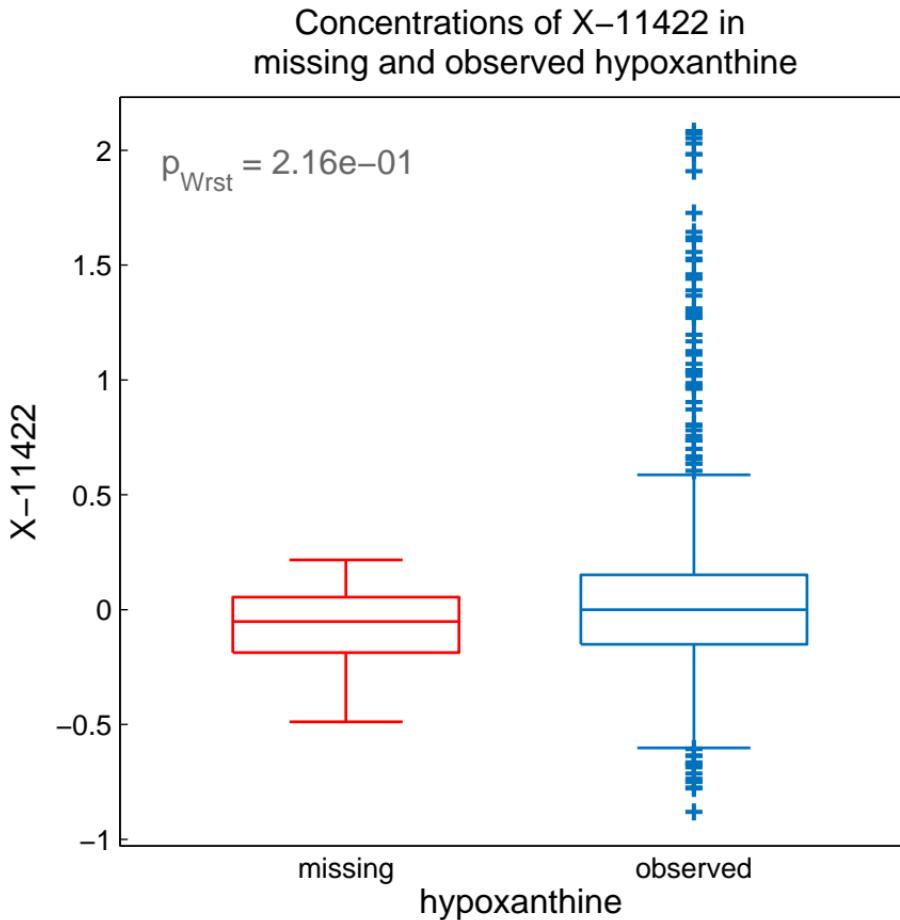
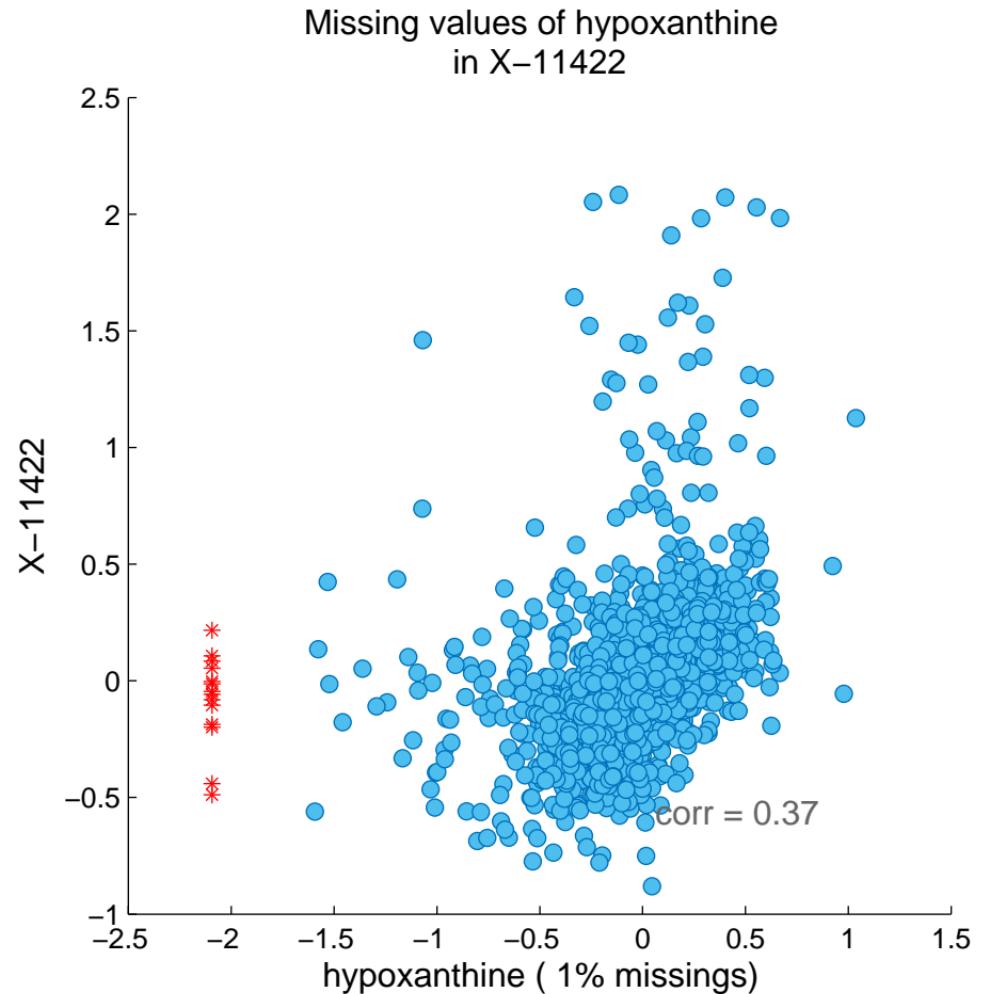
Concentrations of tetradecanedioate in
missing and observed hexadecanedioate

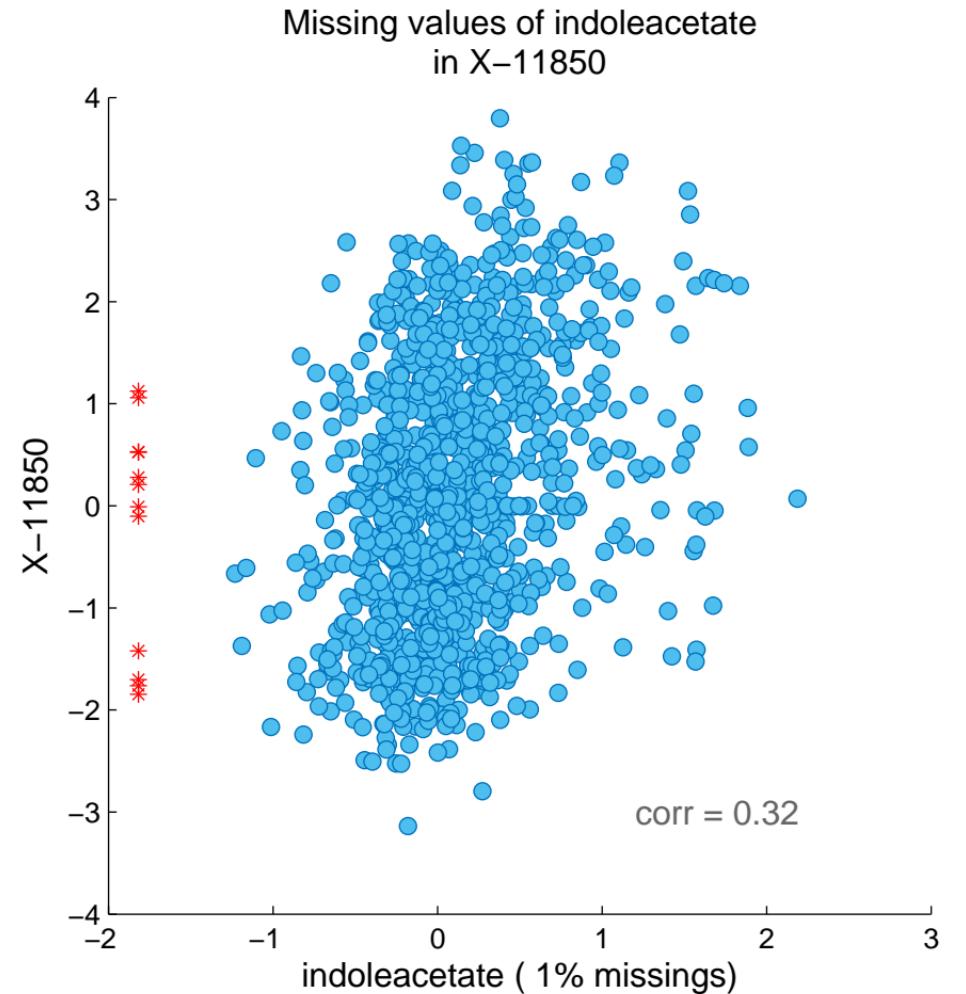




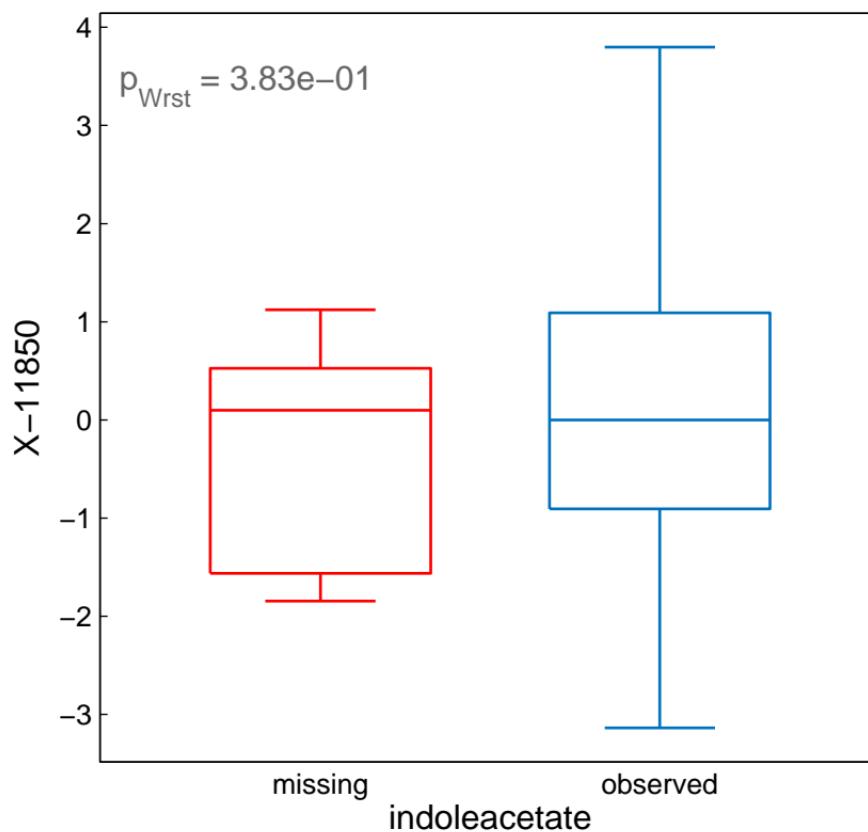


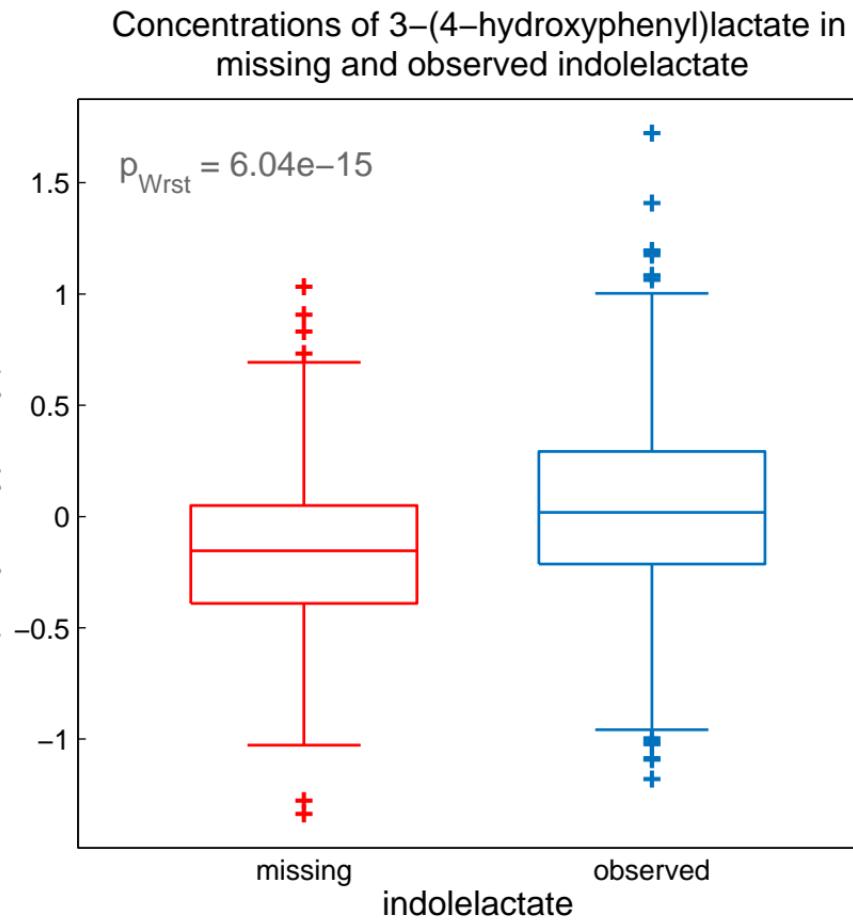
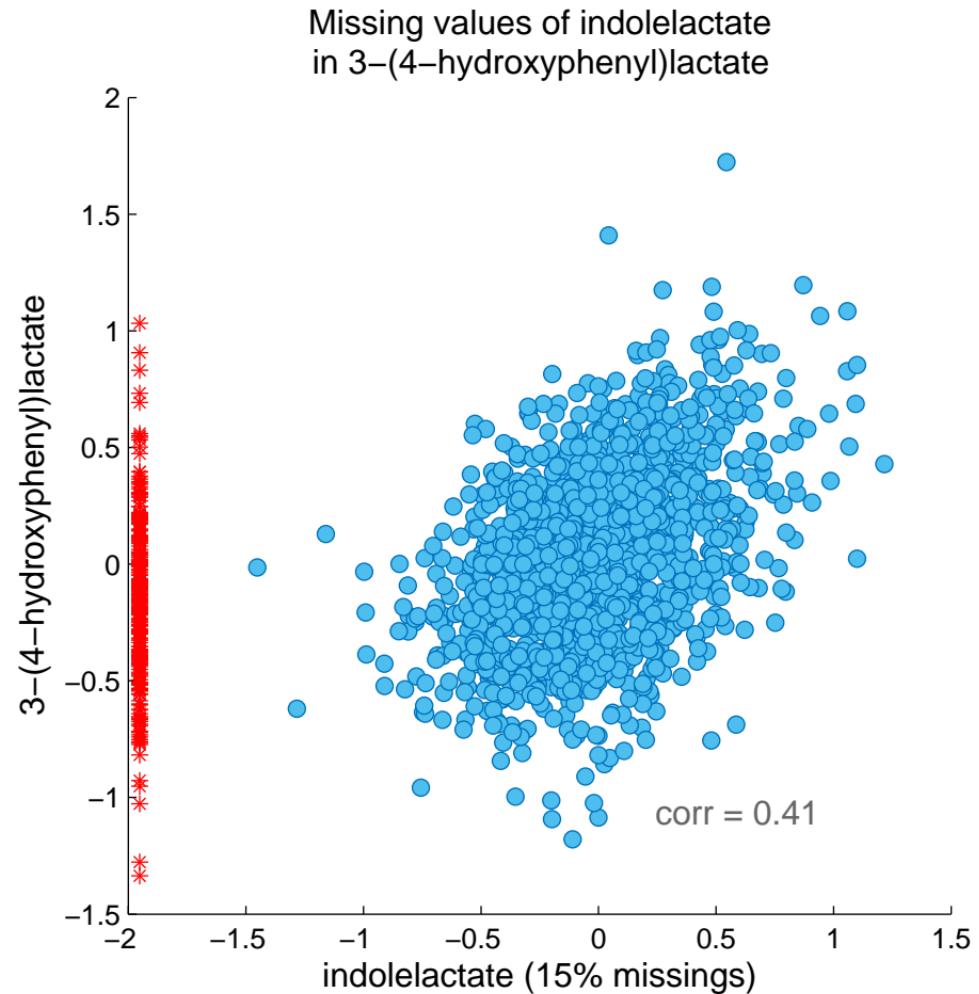


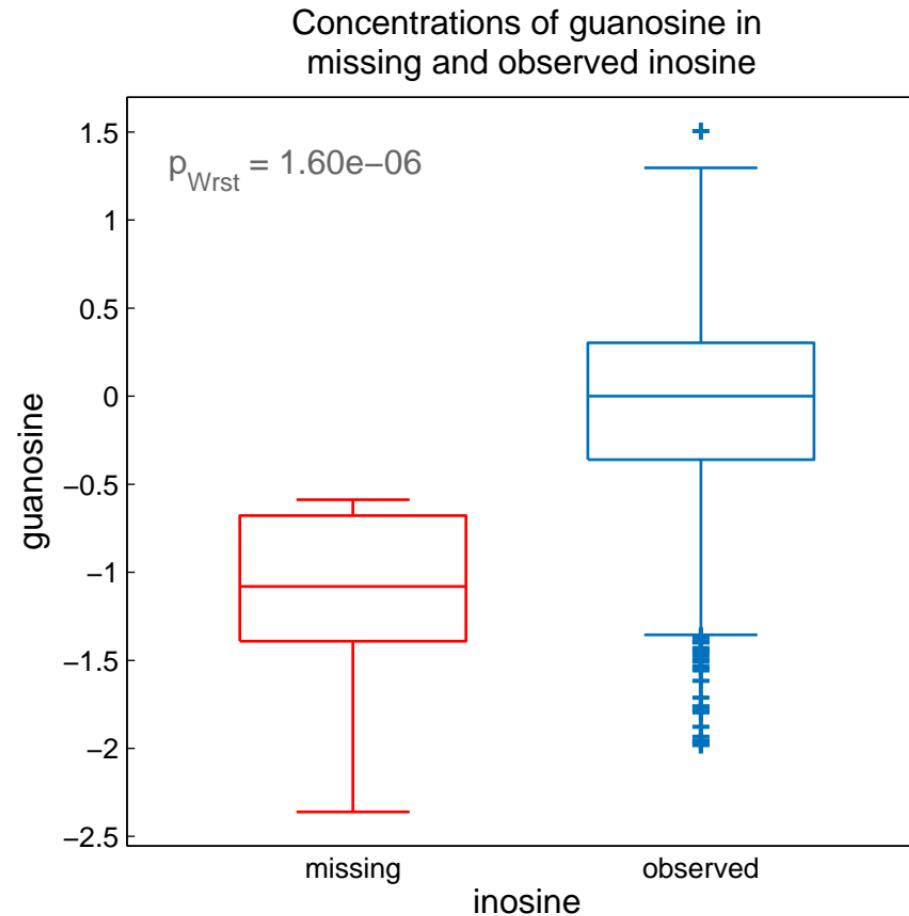
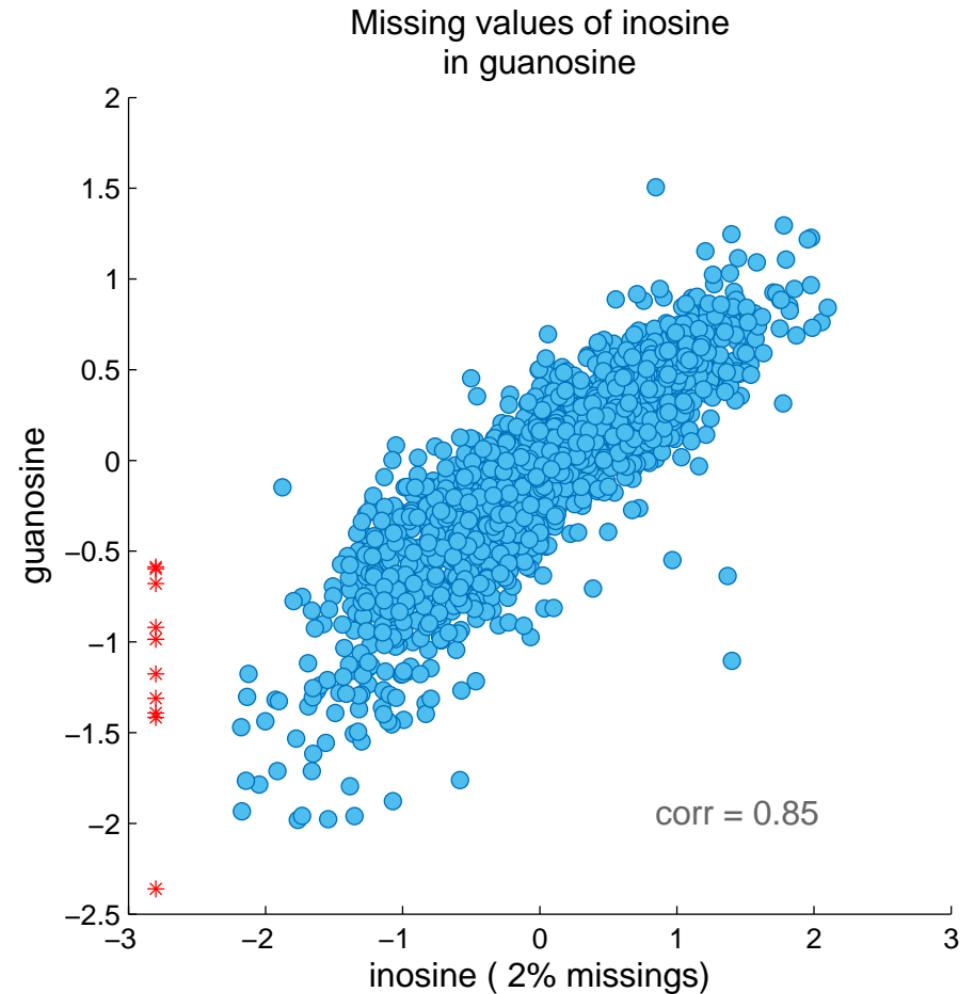


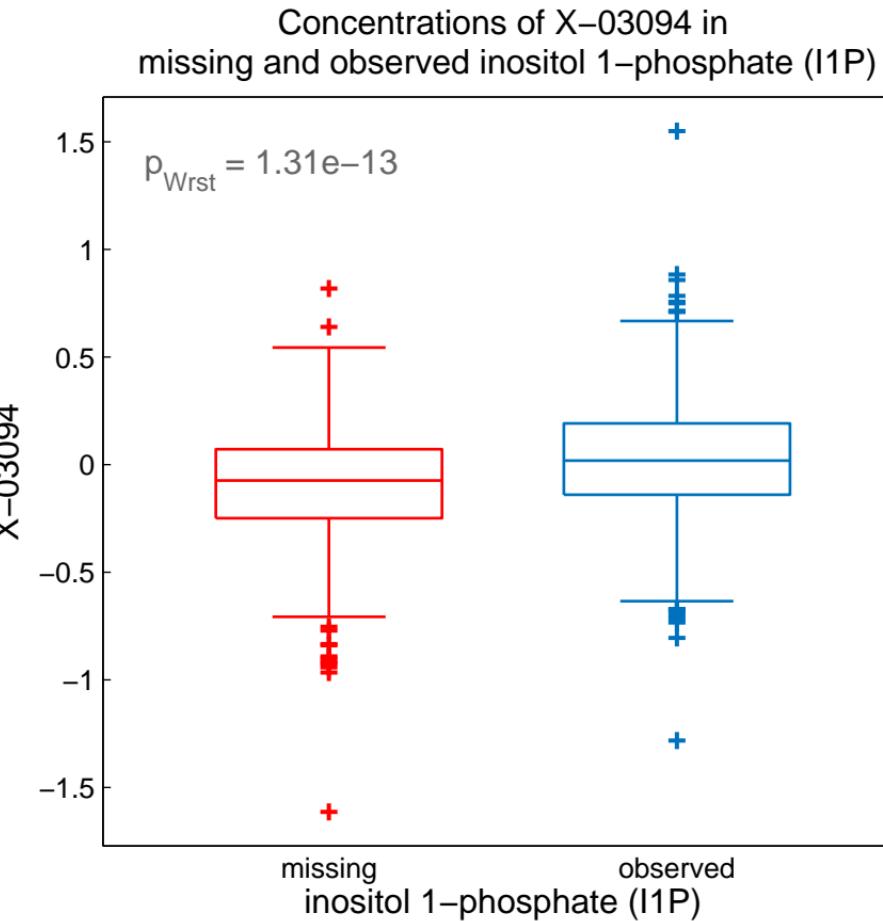
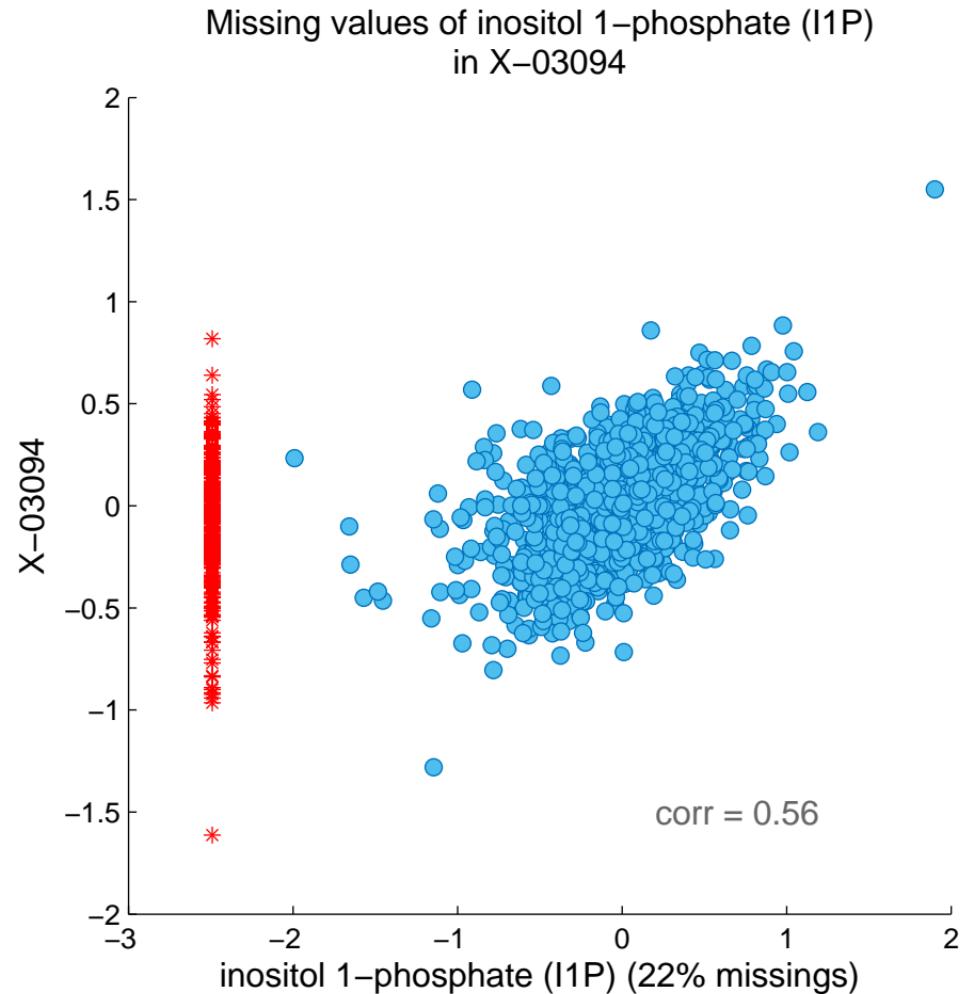


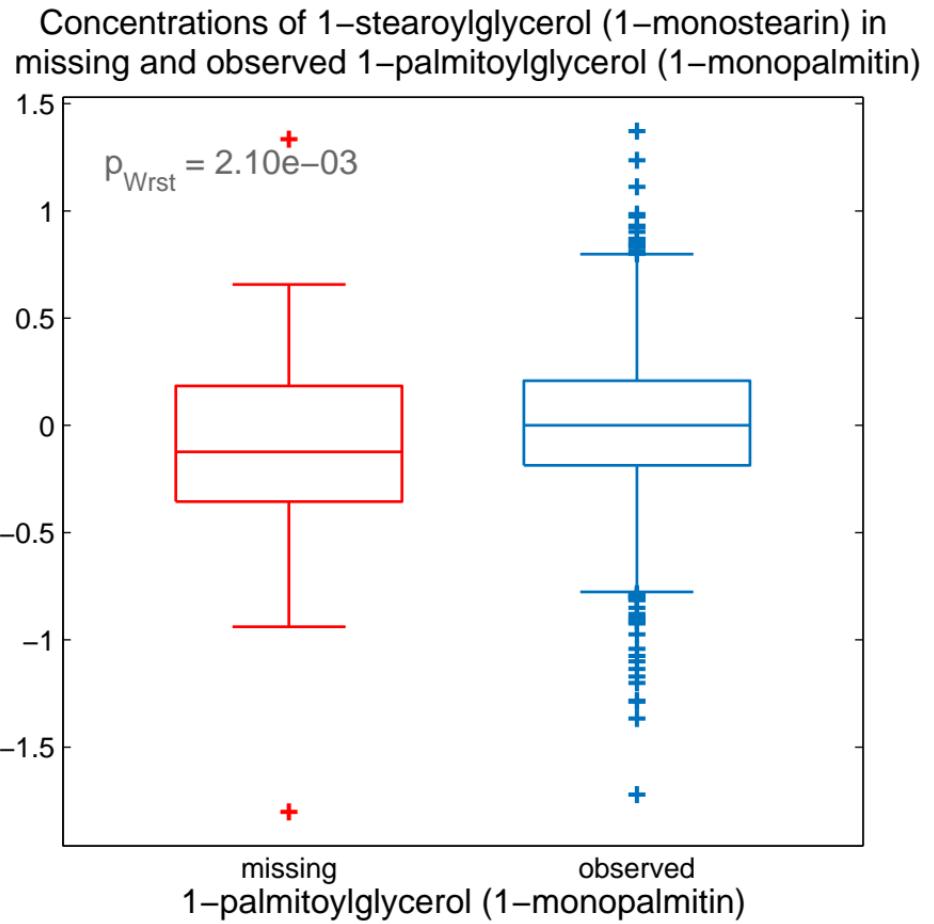
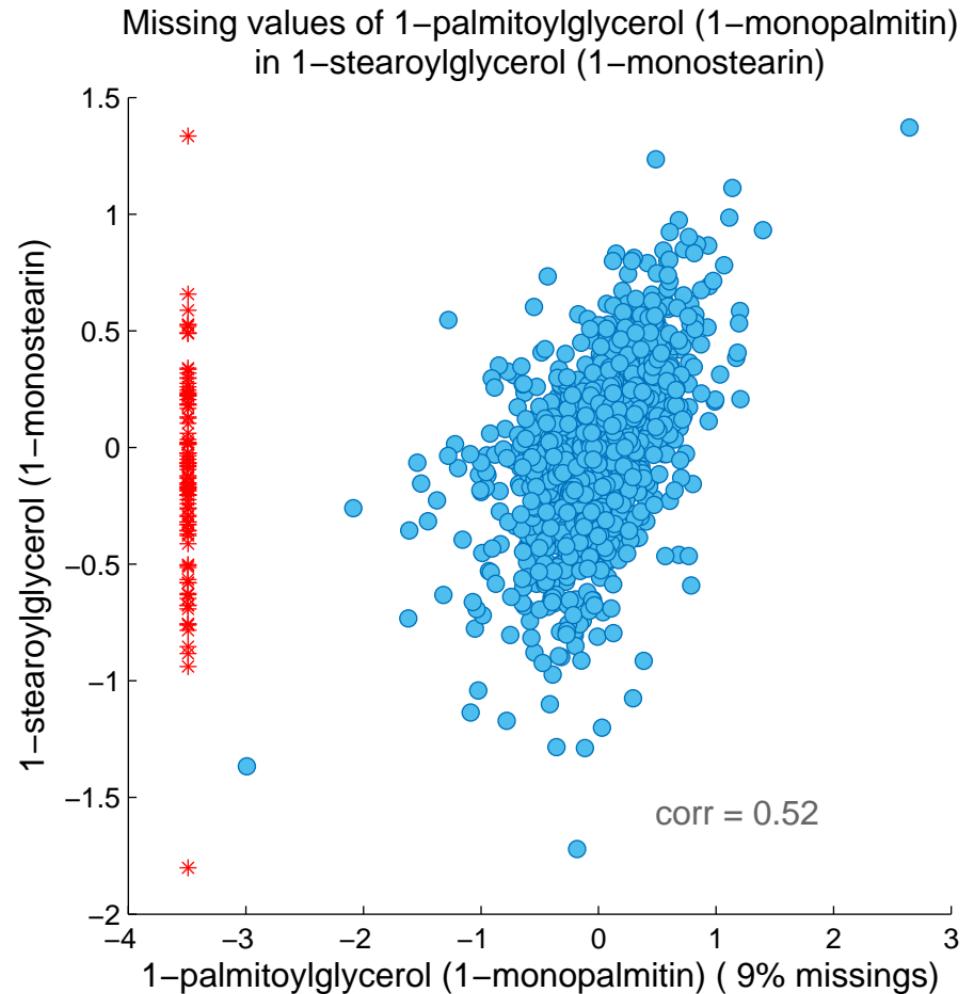
Concentrations of X-11850 in
missing and observed indoleacetate

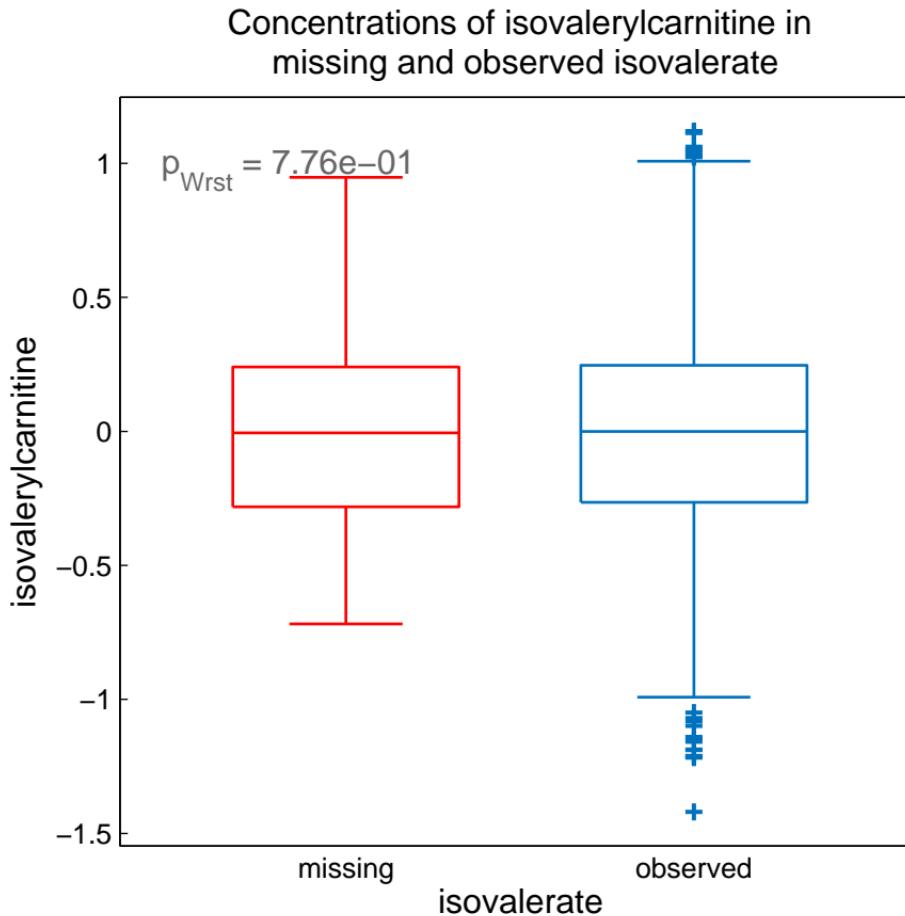
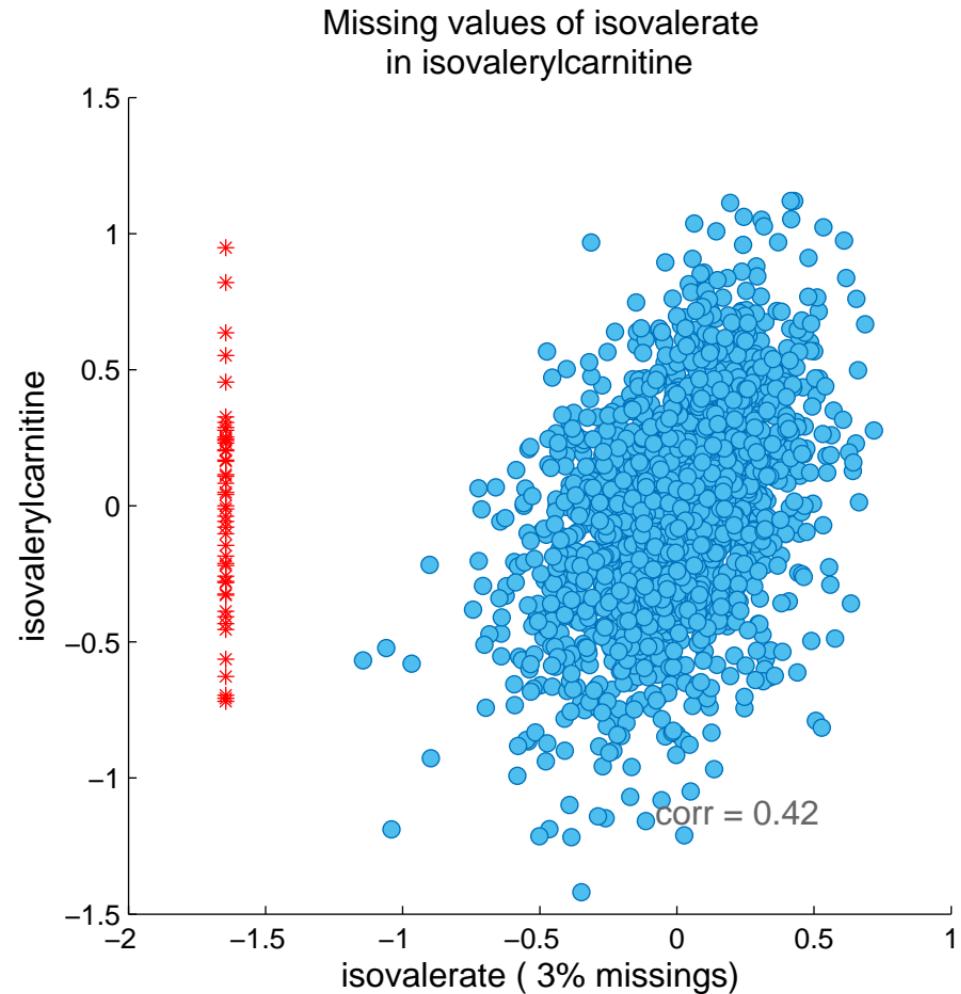


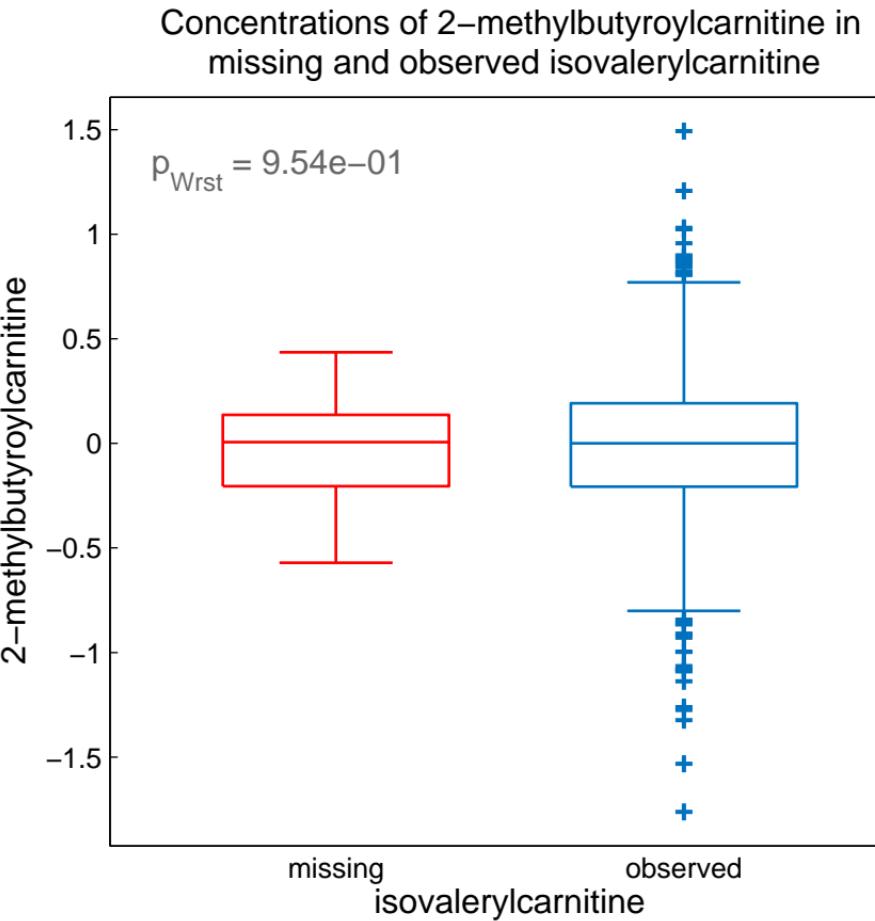
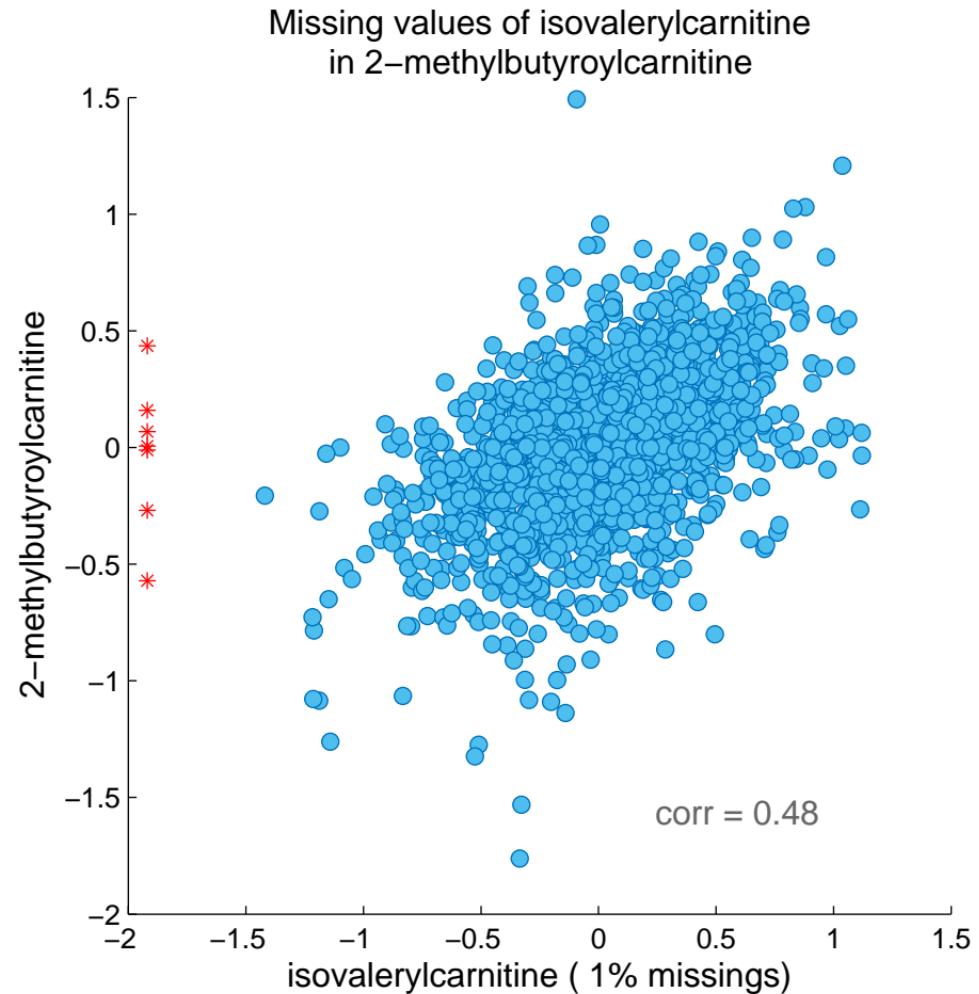


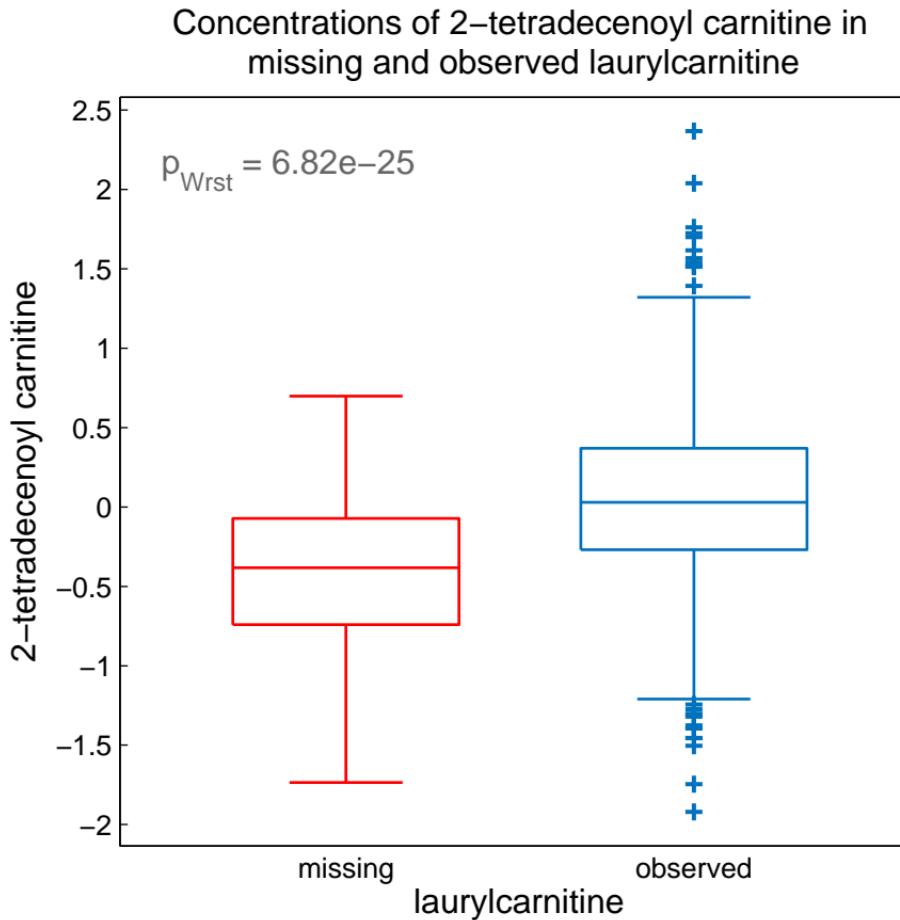
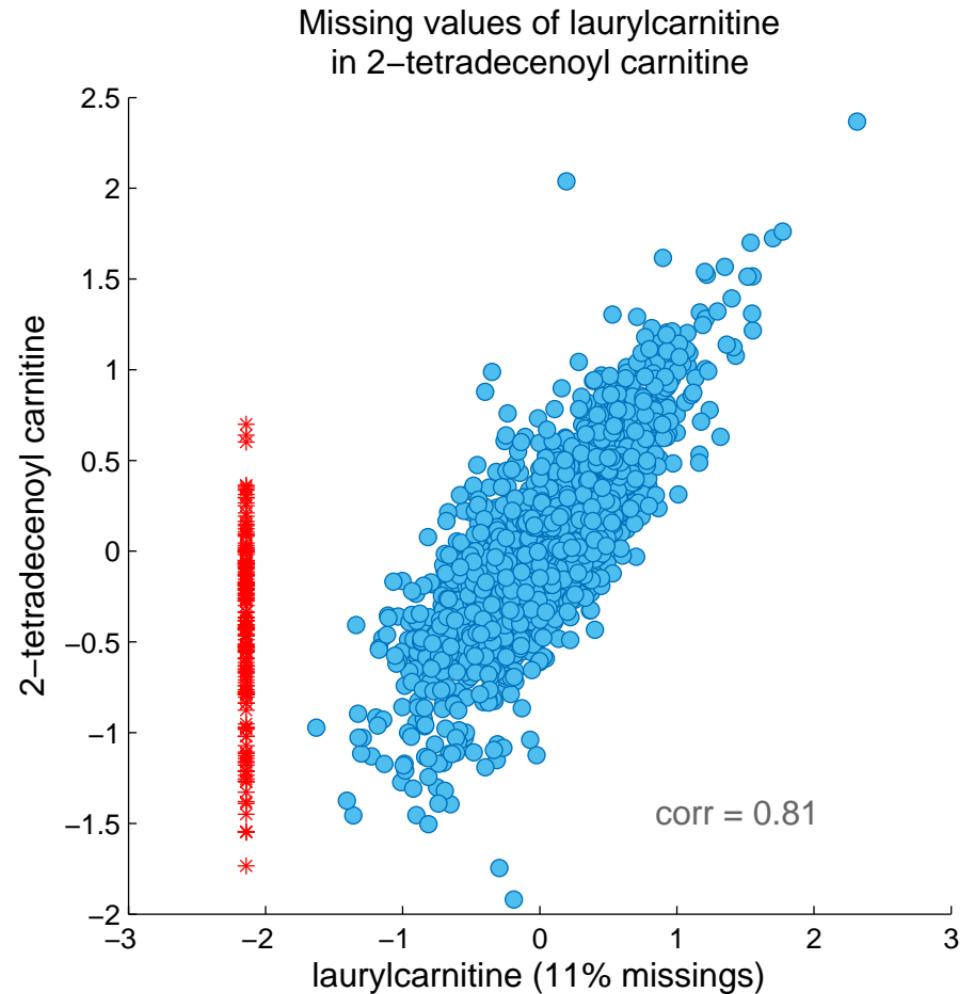


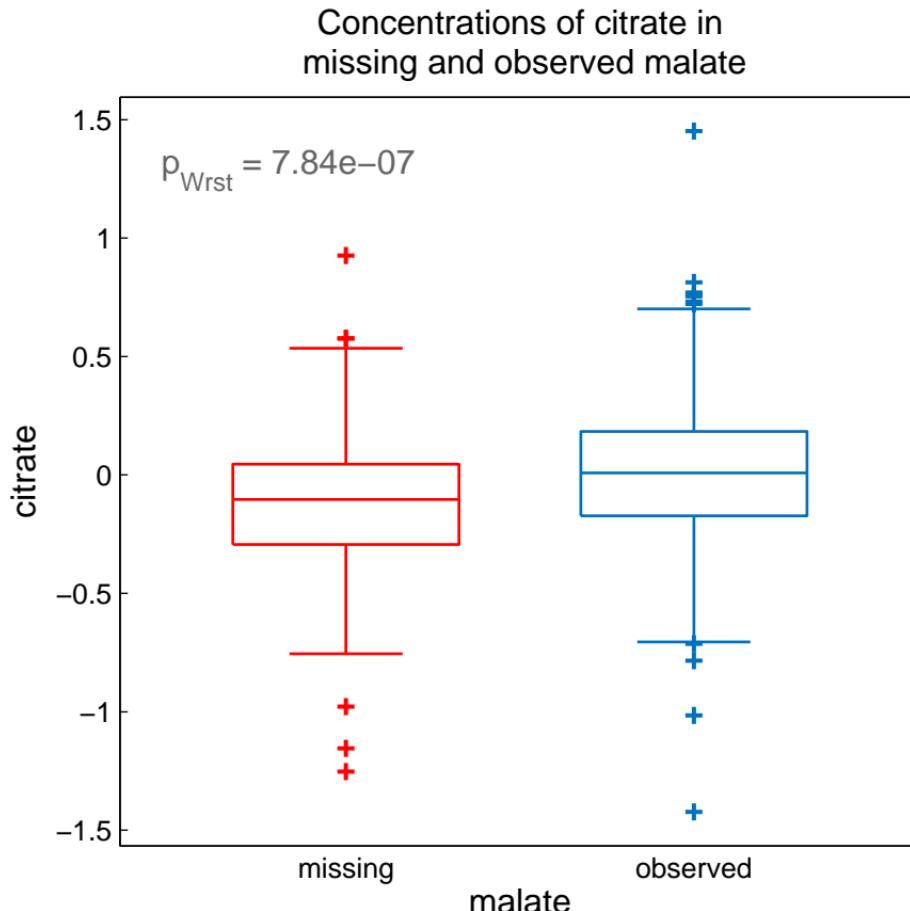
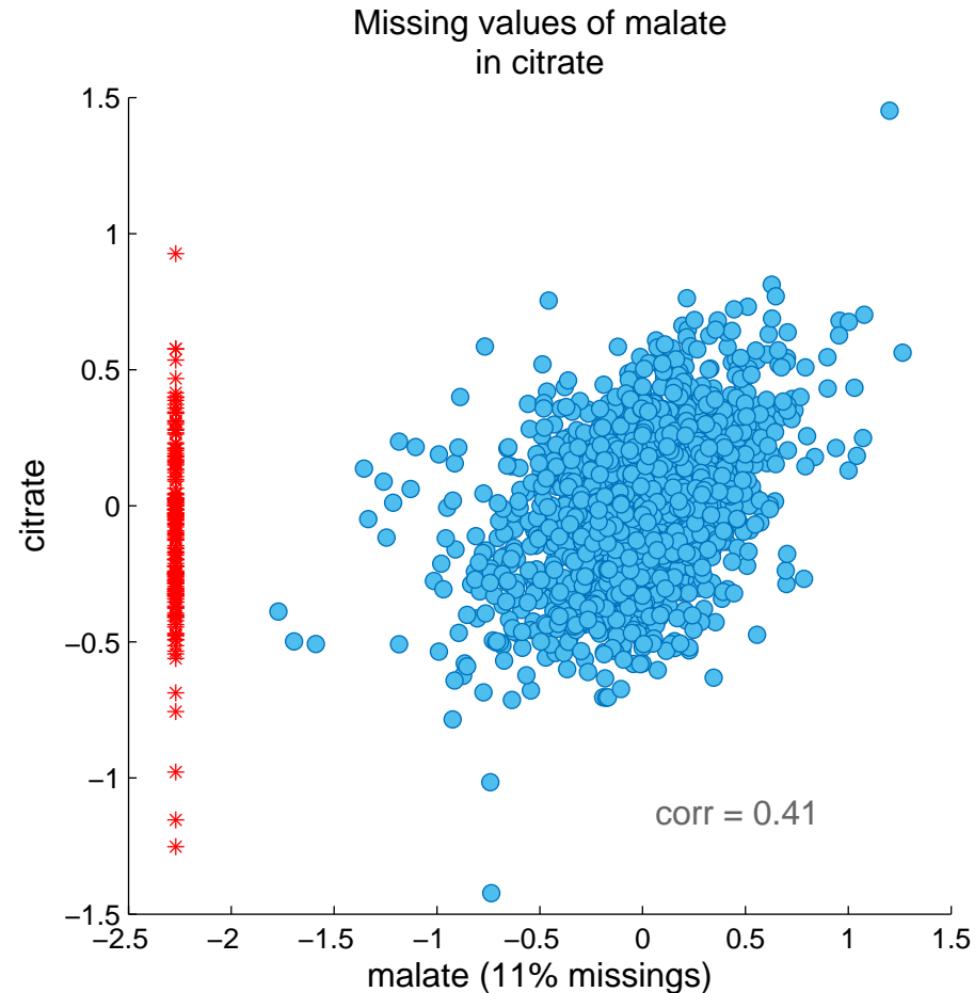


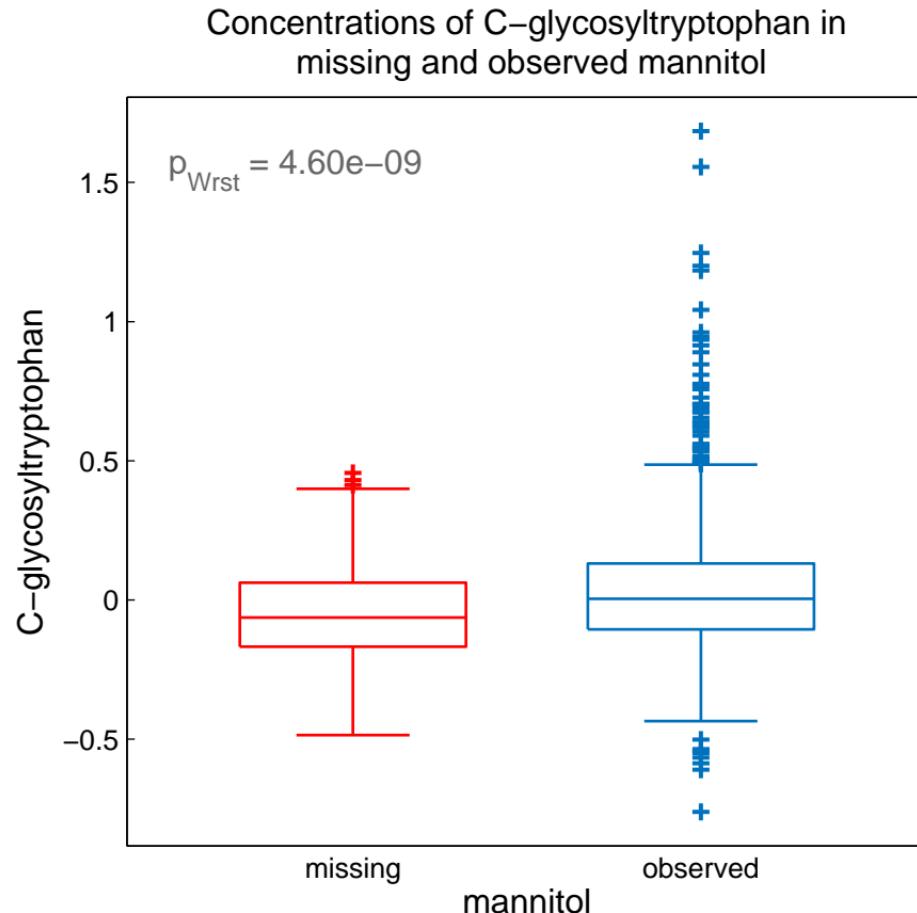
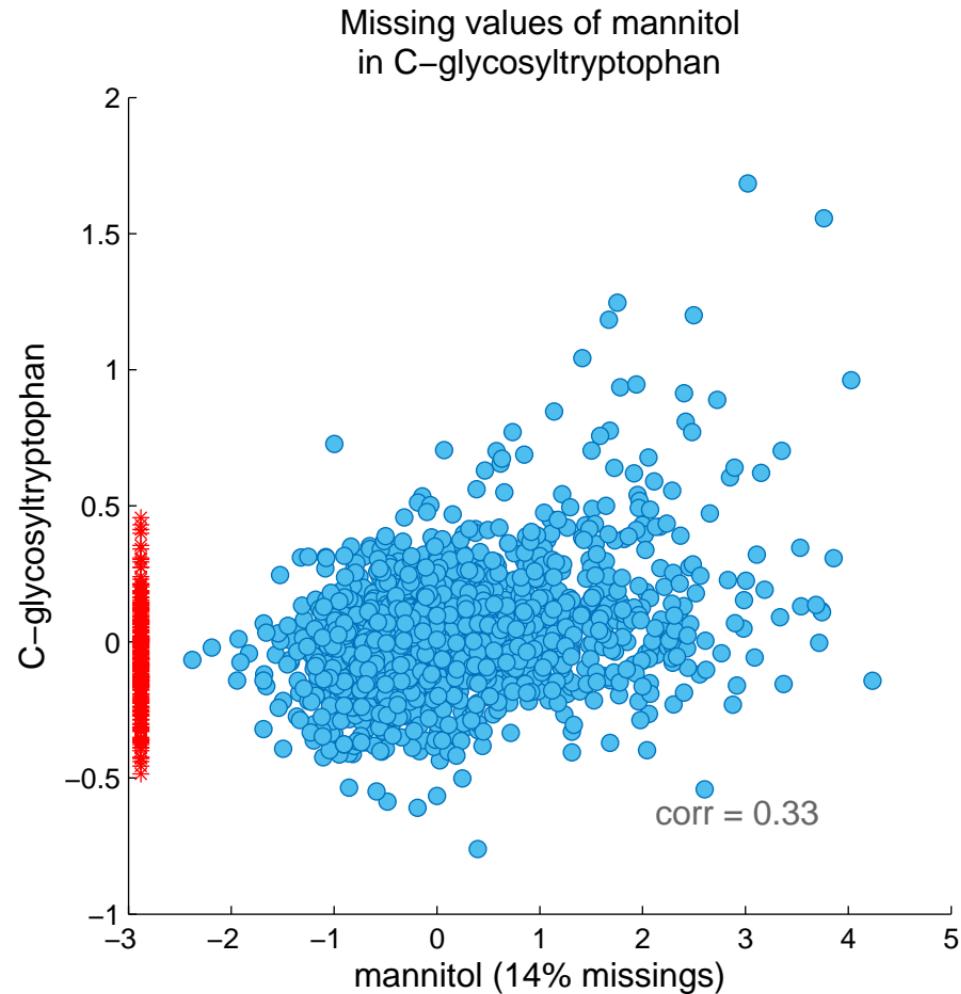




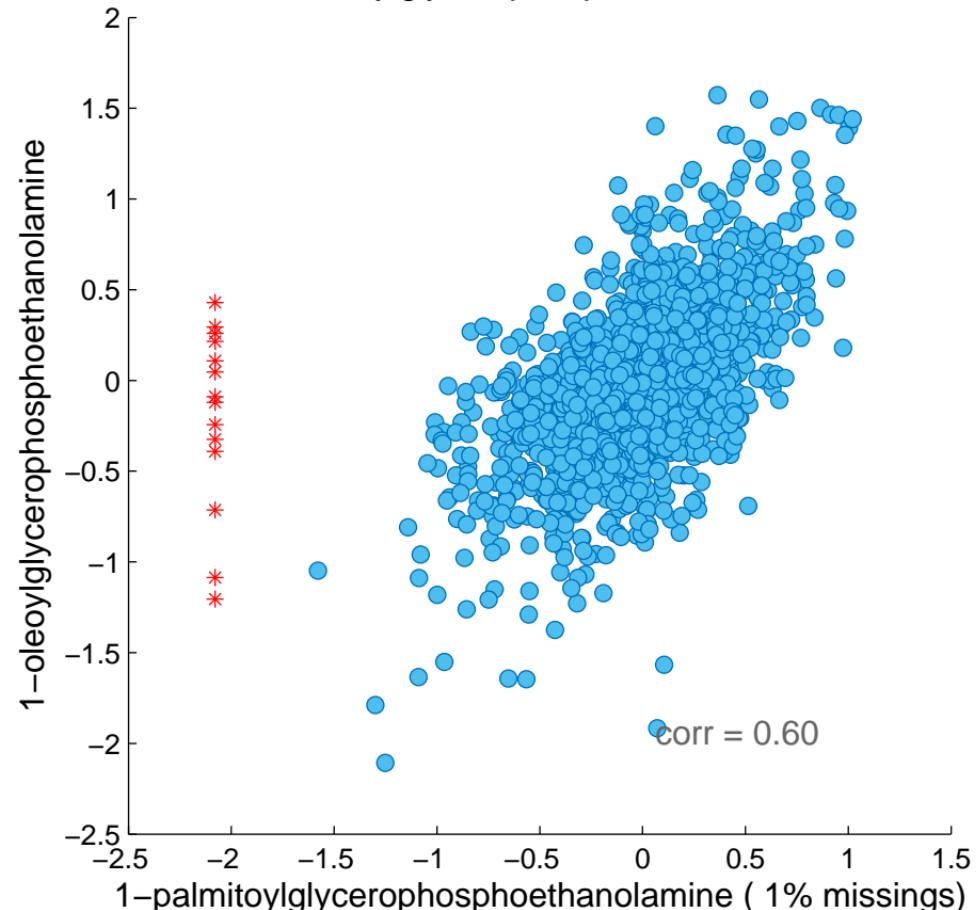




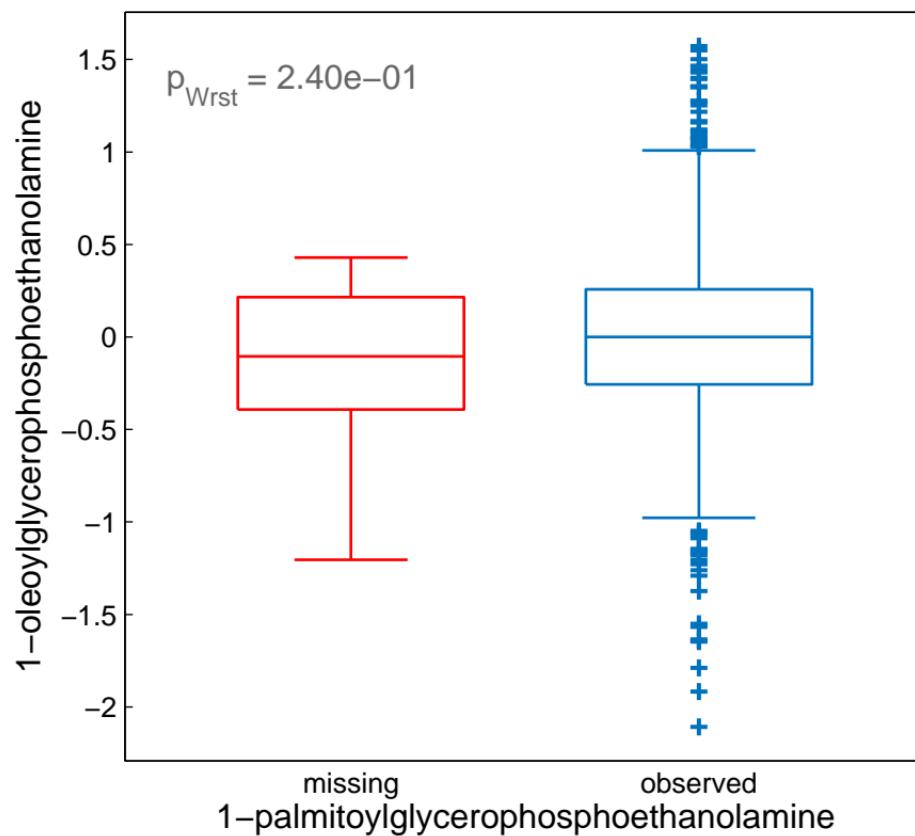


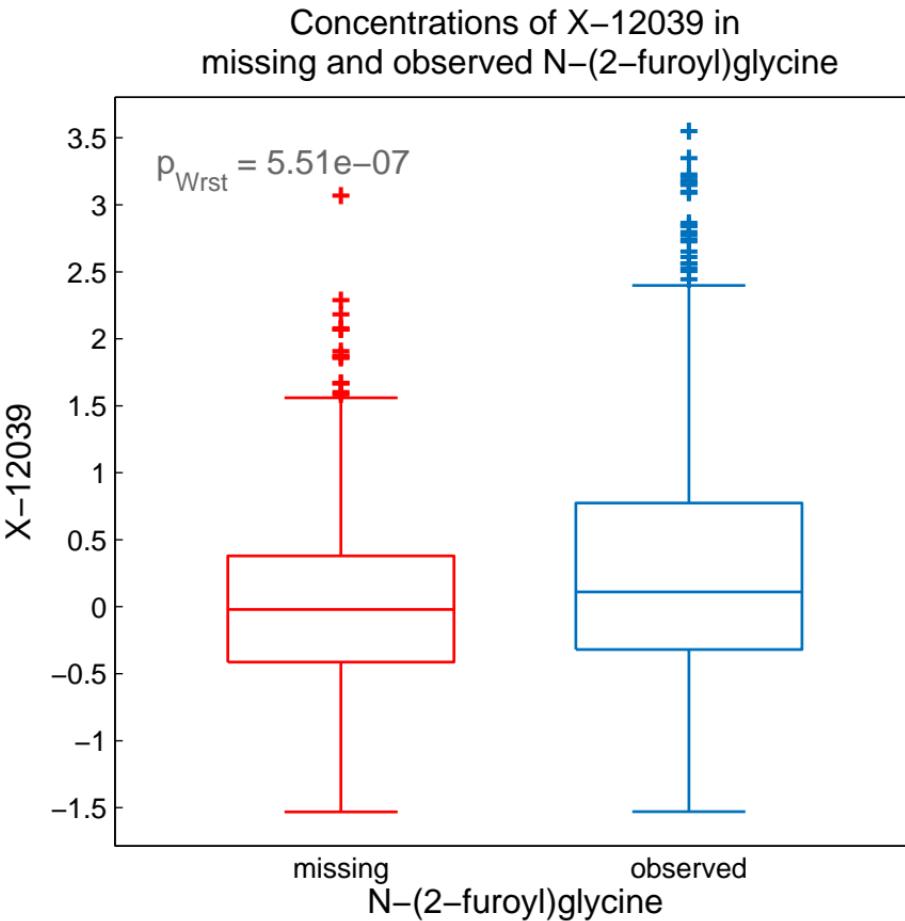
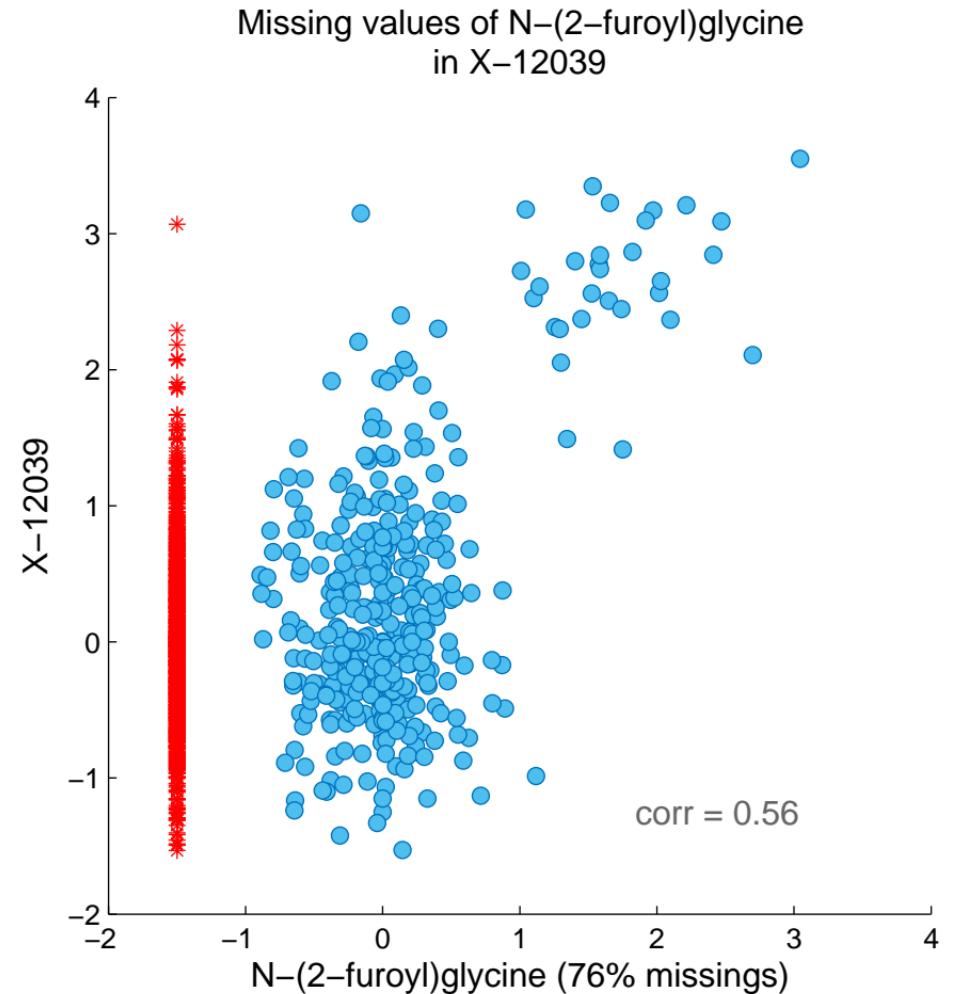


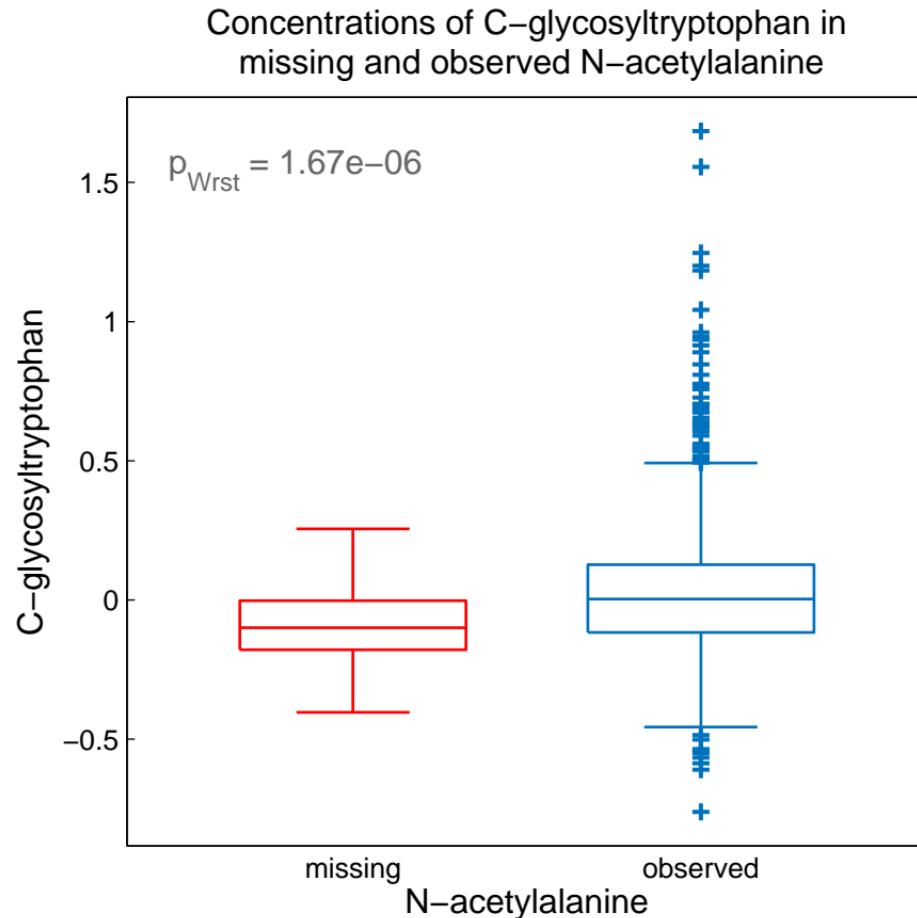
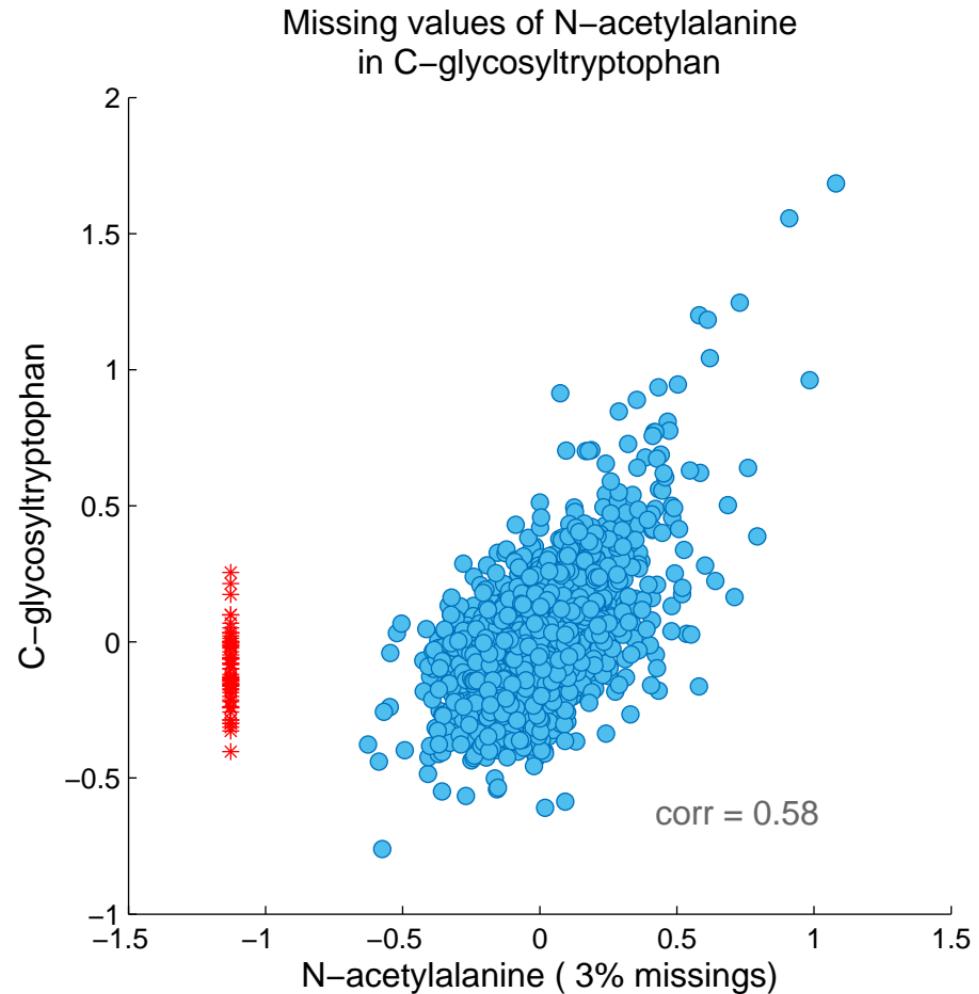
Missing values of 1-palmitoylglycerophosphoethanolamine
in 1-oleoylglycerophosphoethanolamine

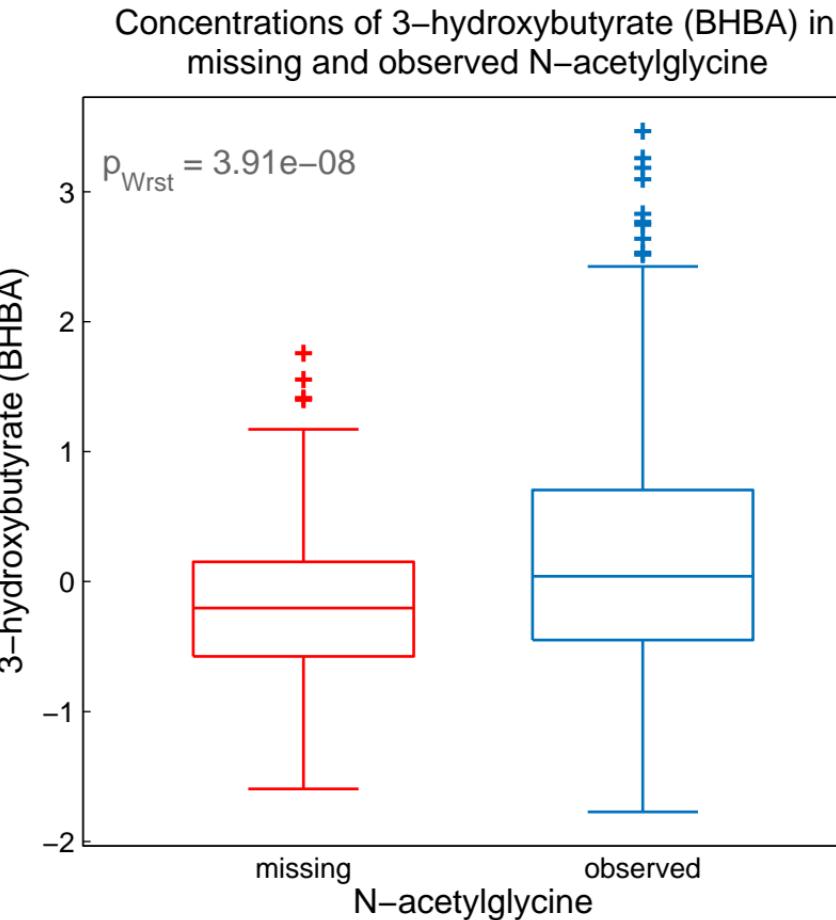
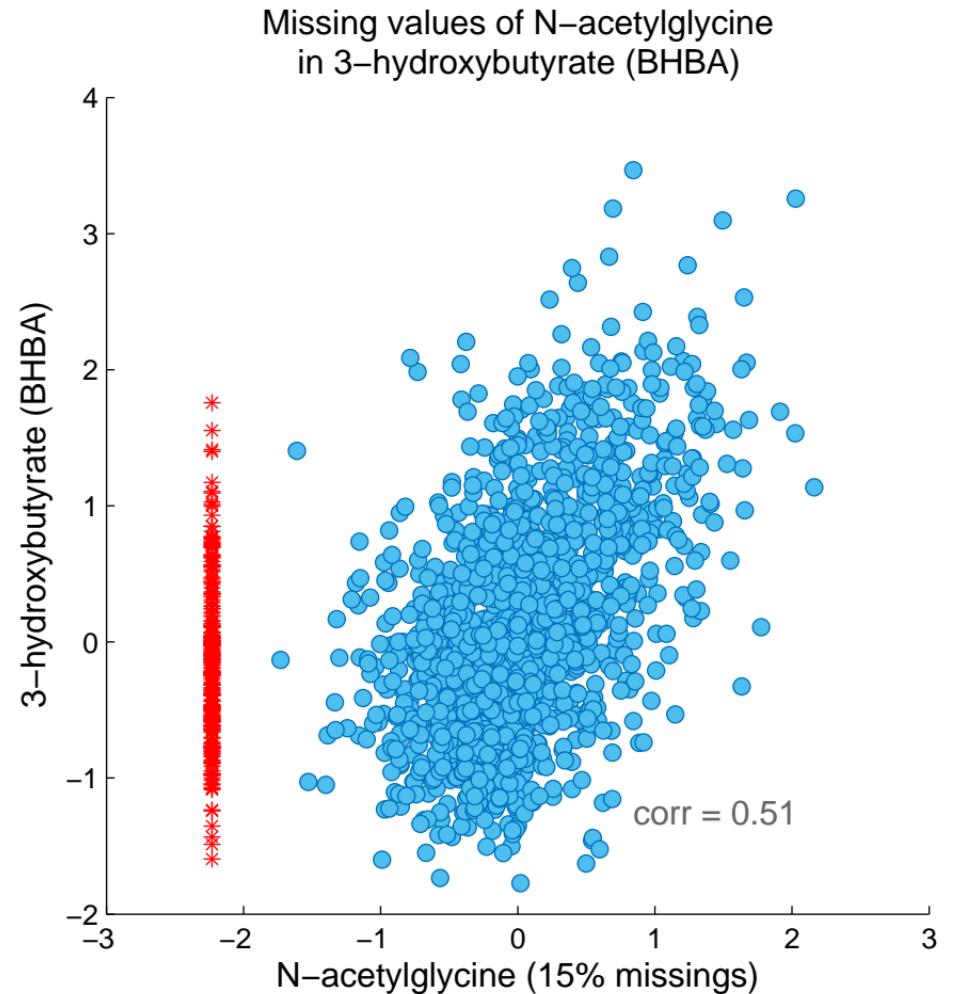


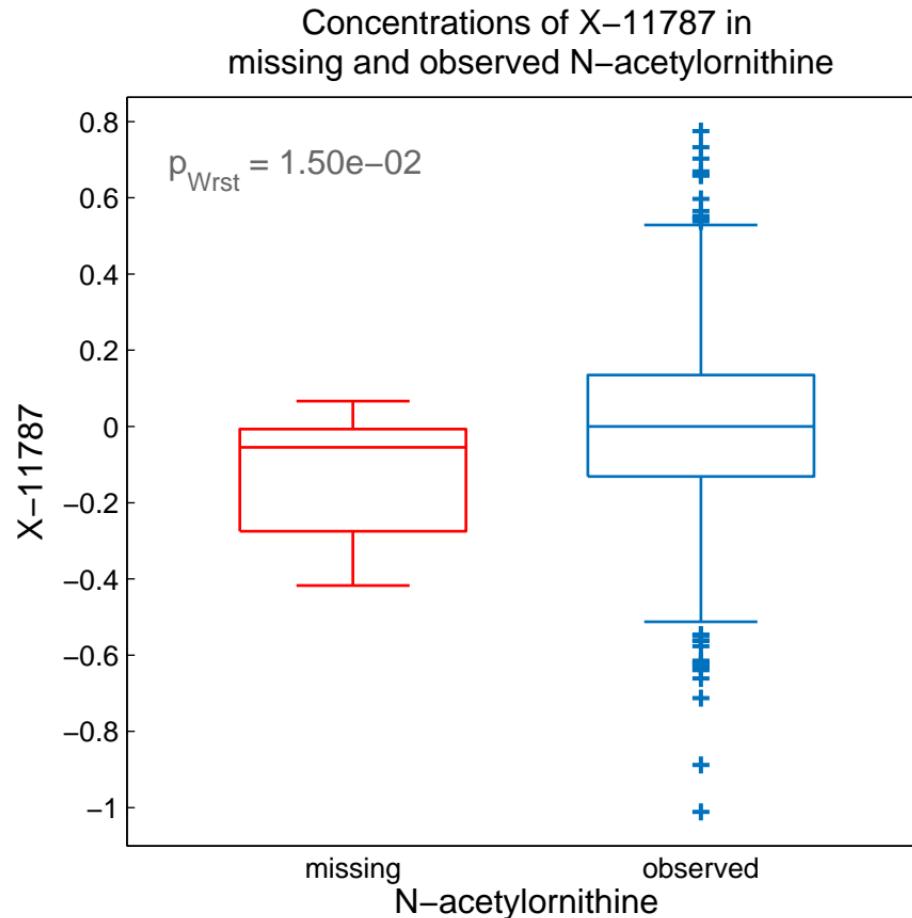
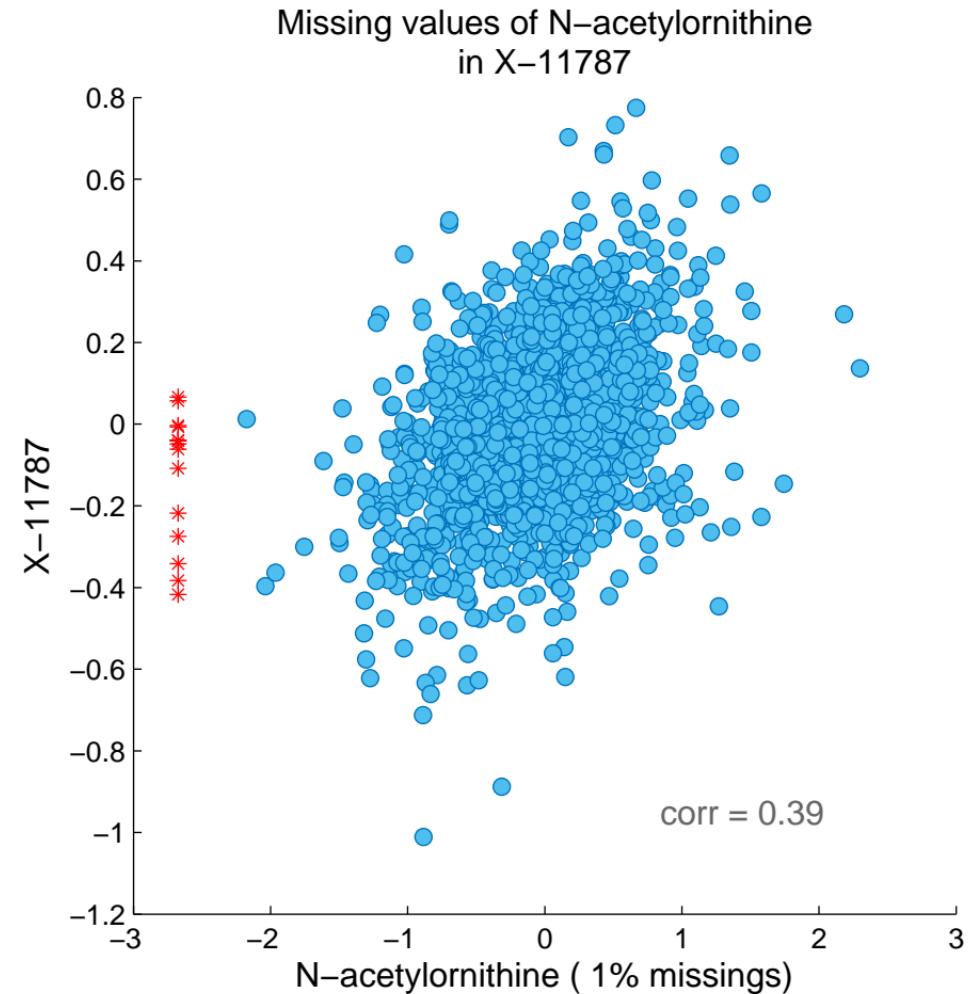
Concentrations of 1-oleoylglycerophosphoethanolamine in
missing and observed 1-palmitoylglycerophosphoethanolamine

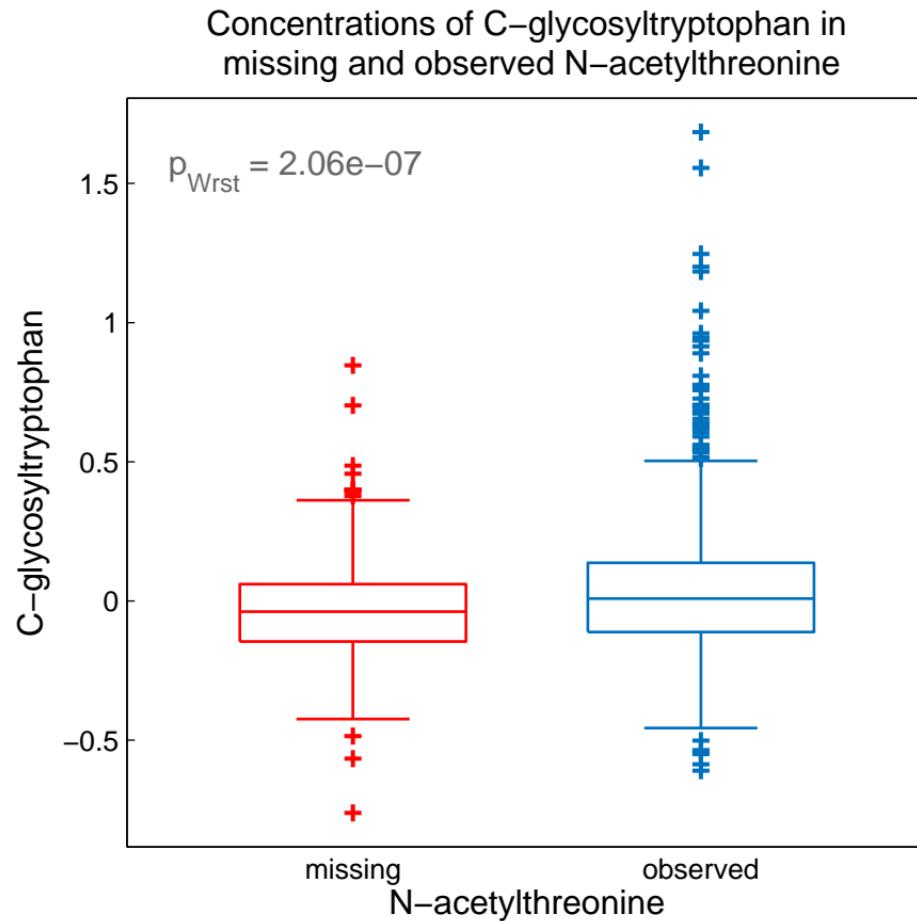
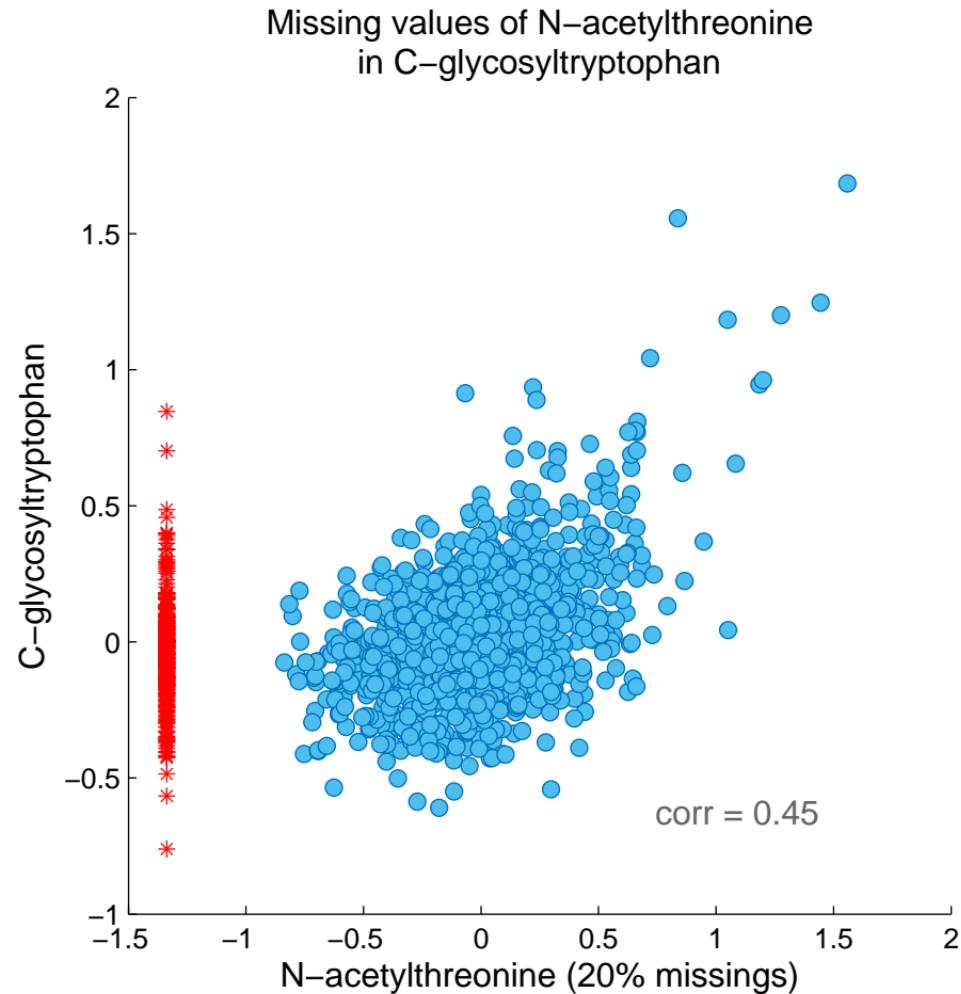


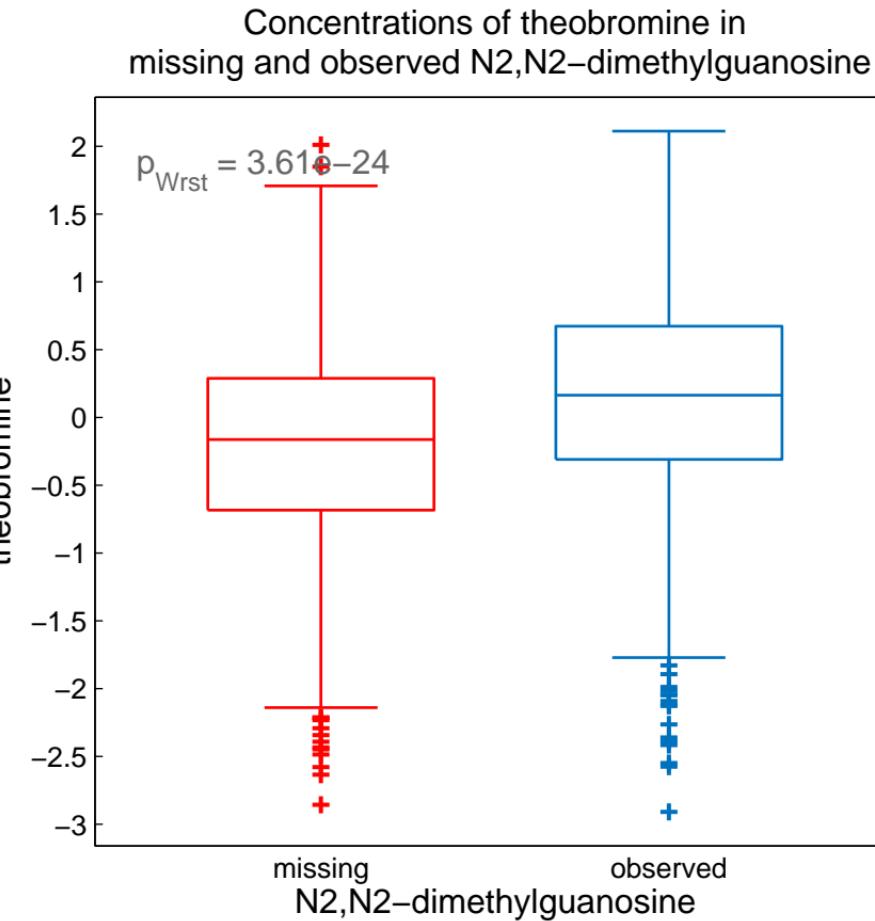
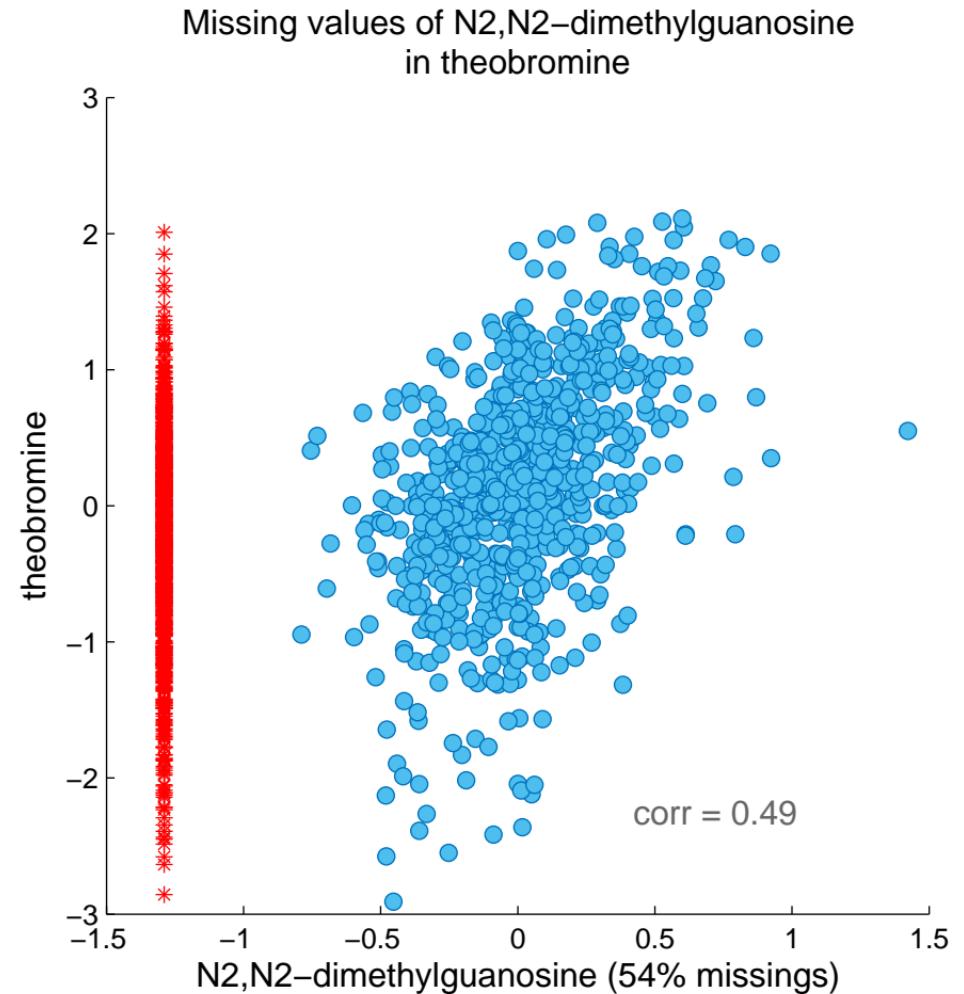


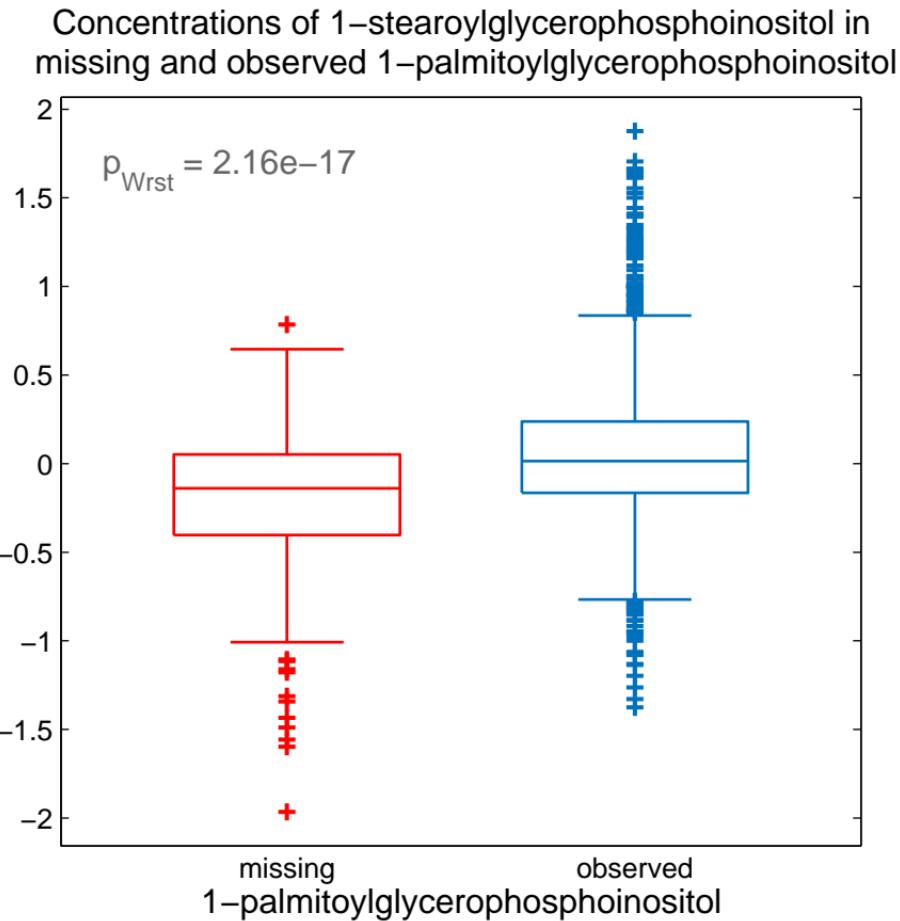
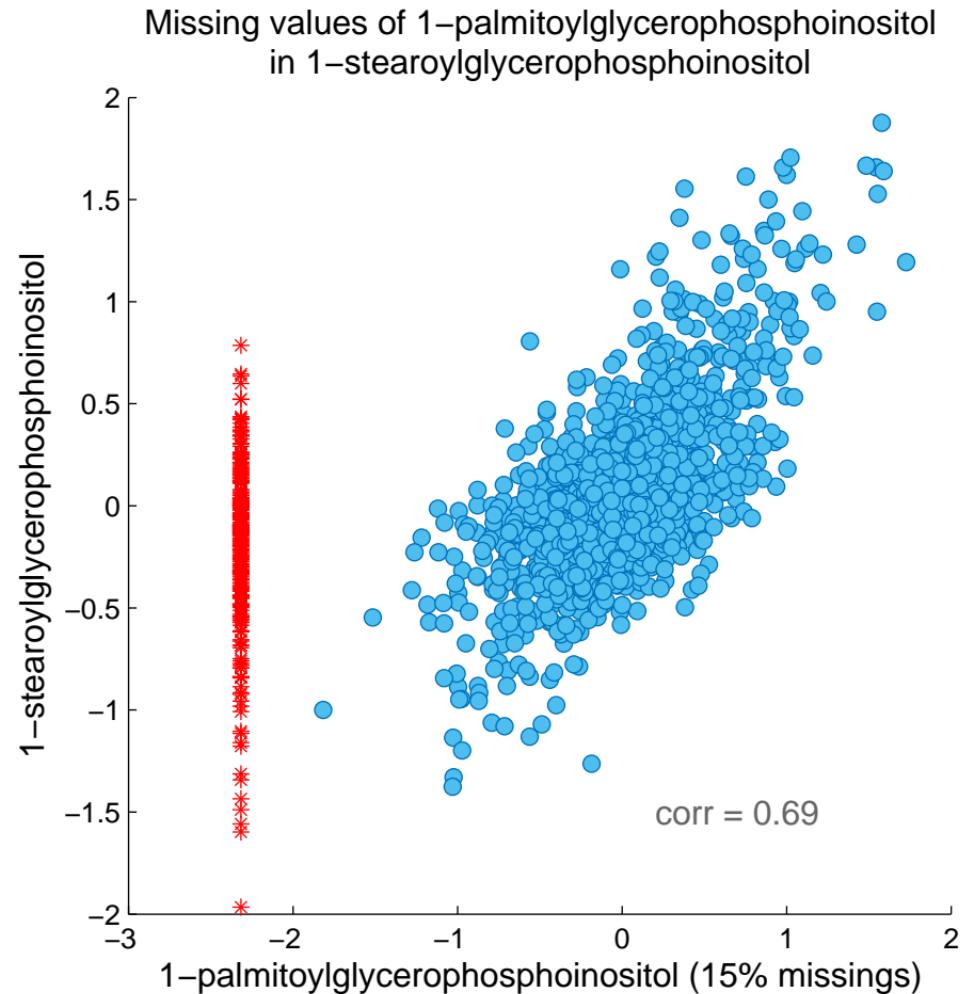


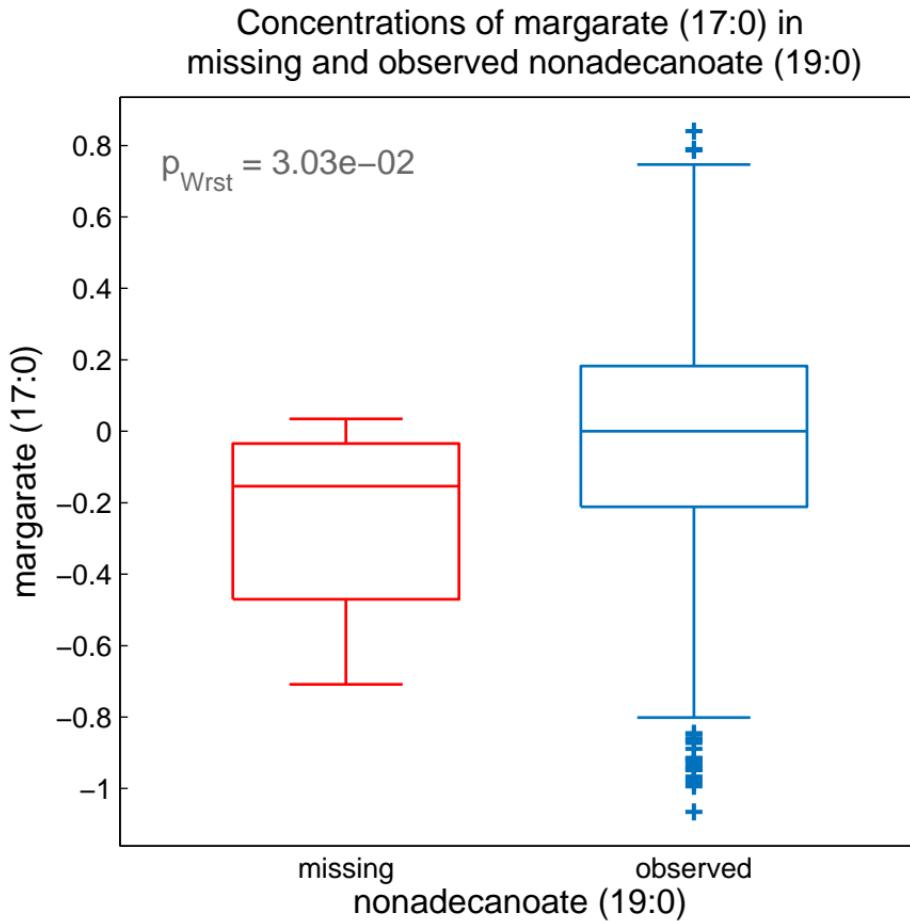
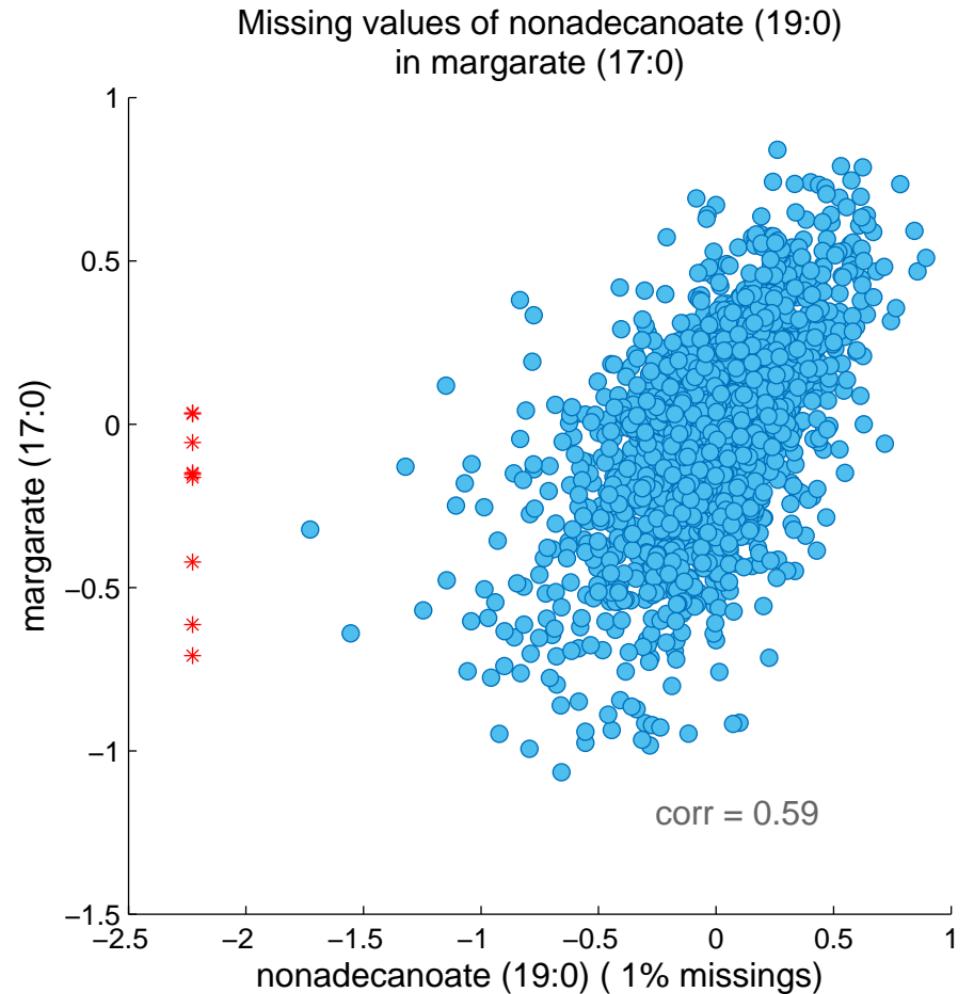


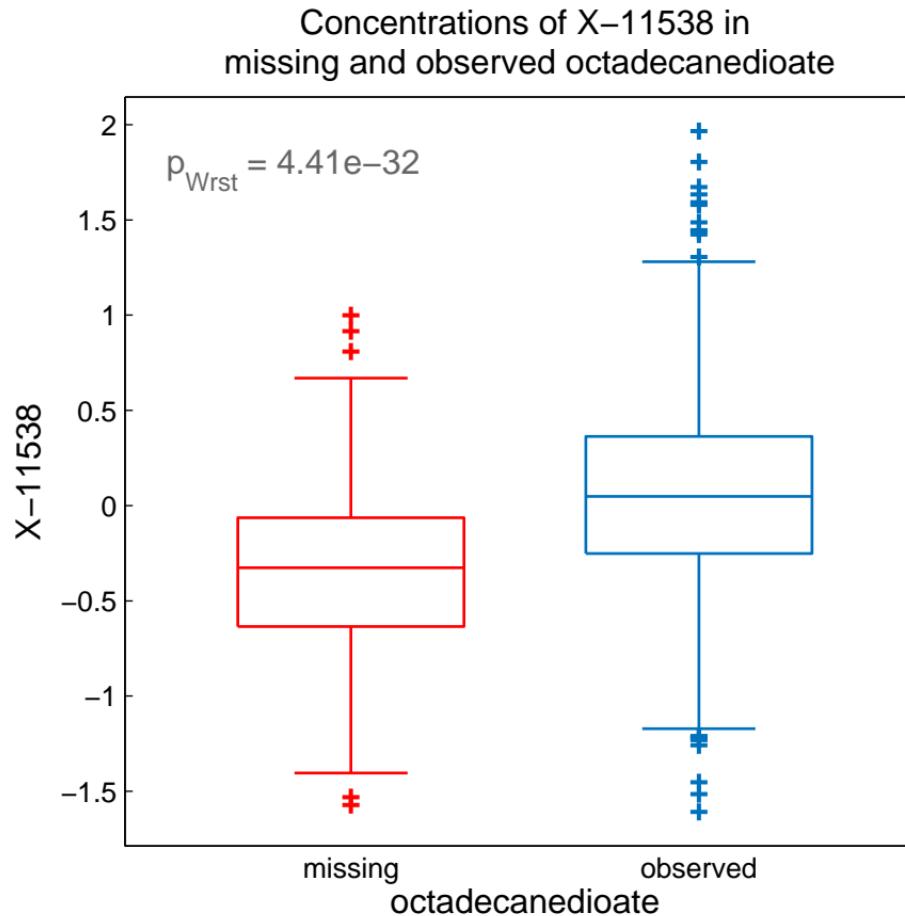
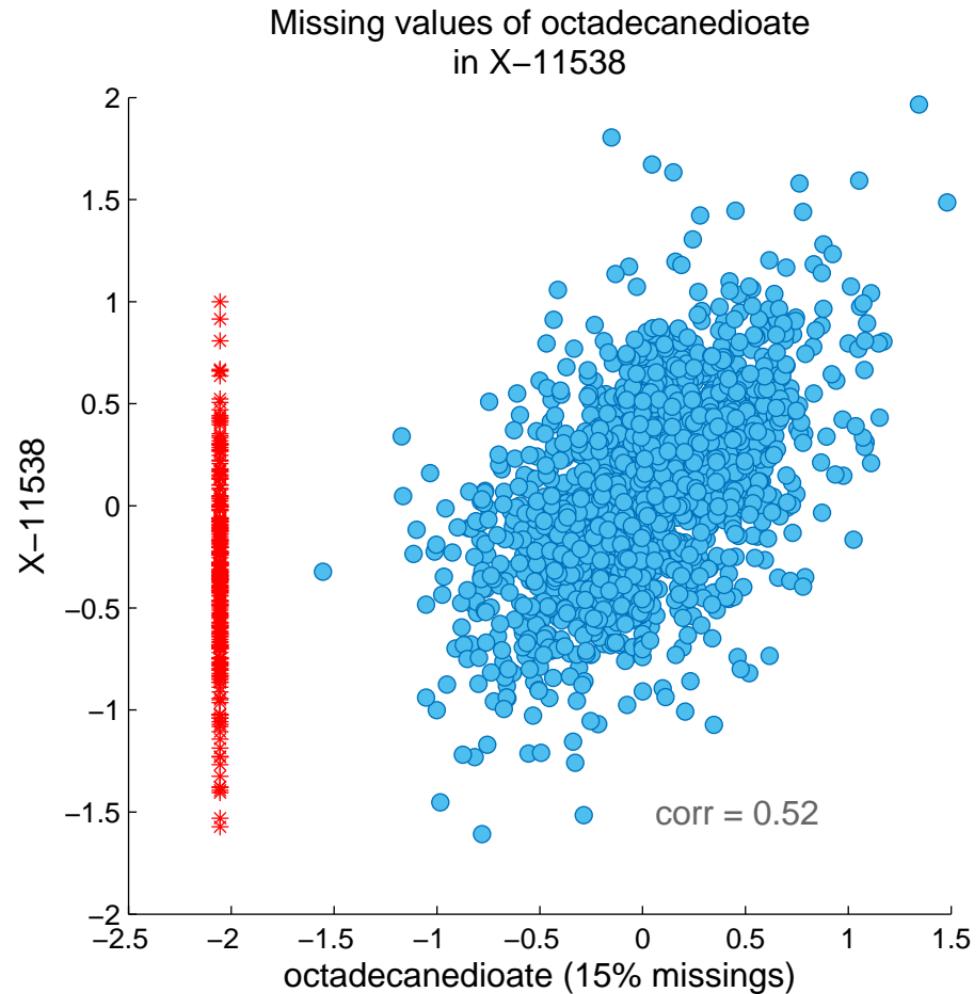


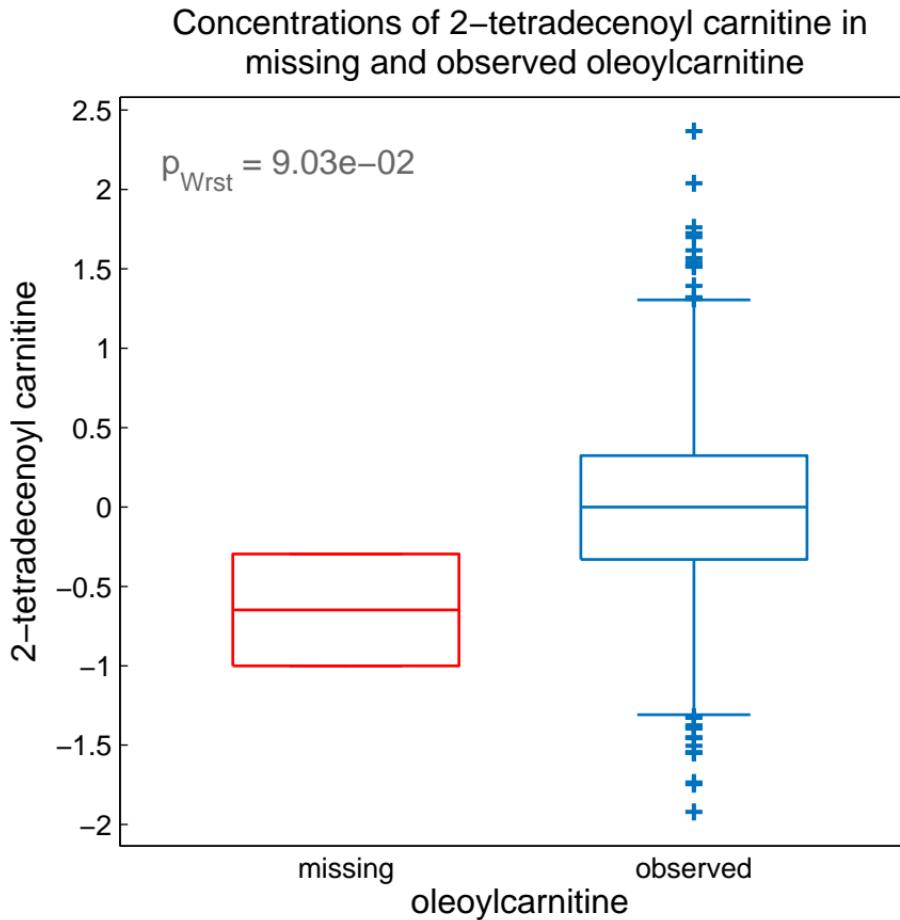
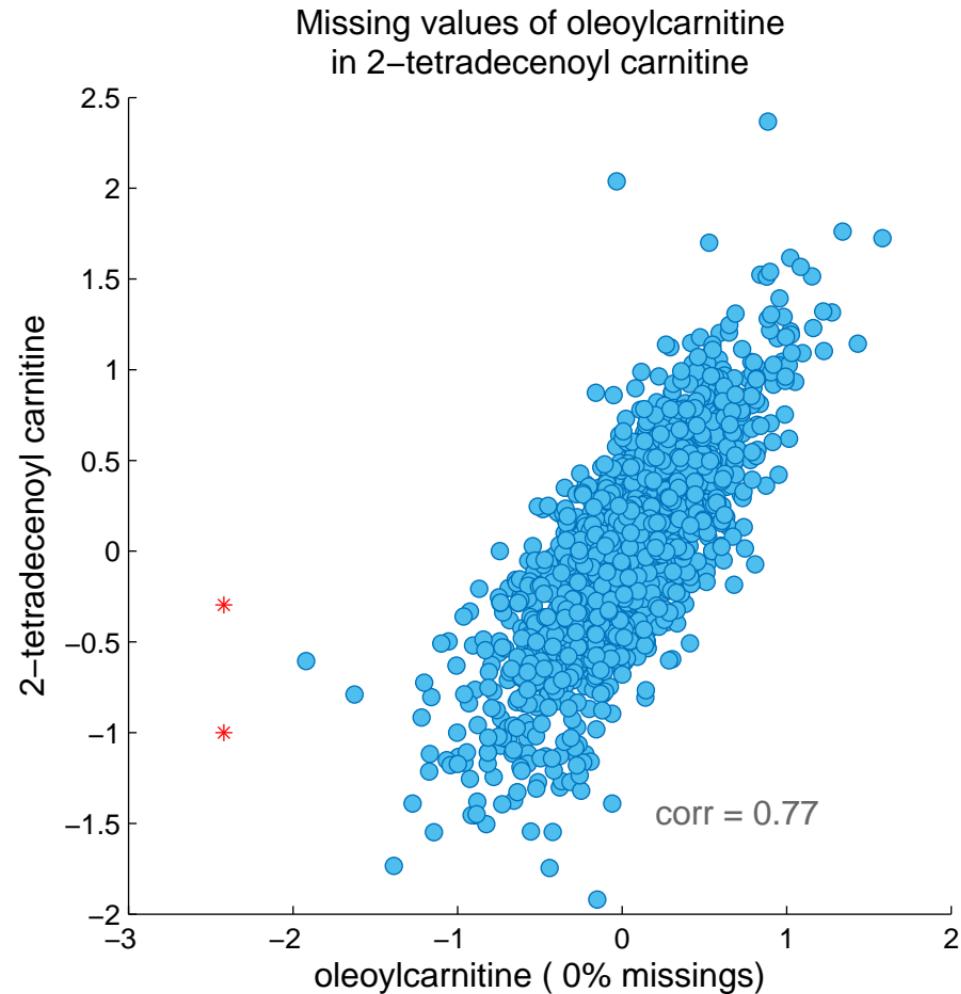


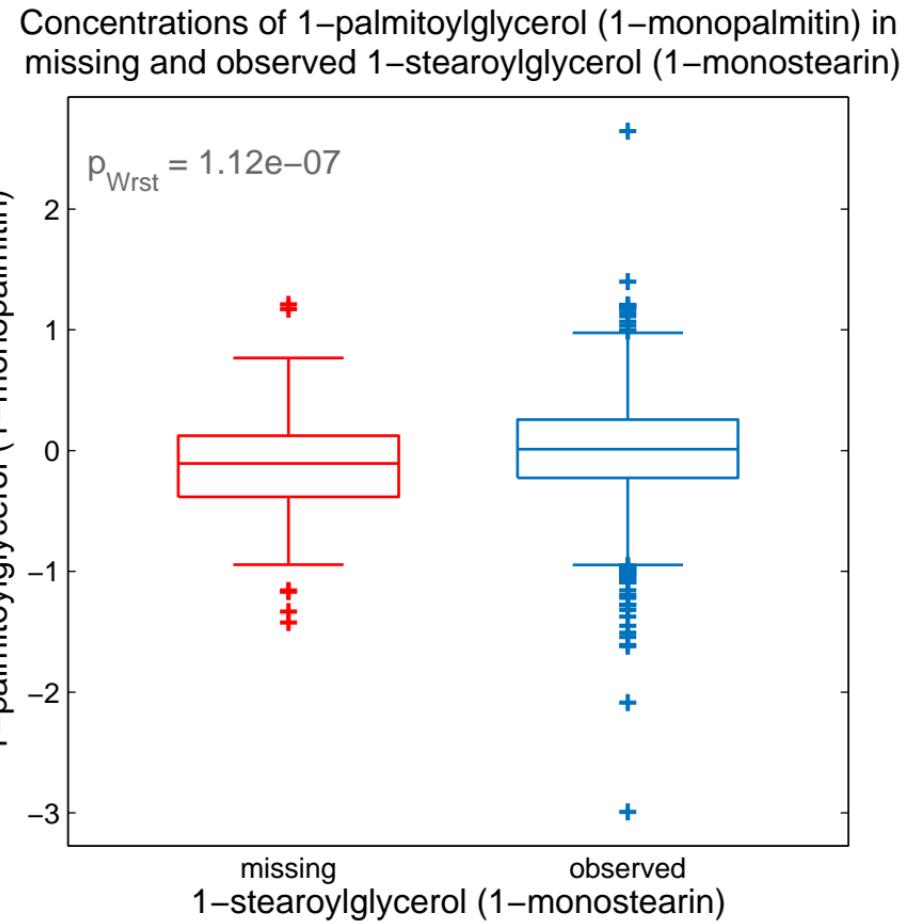
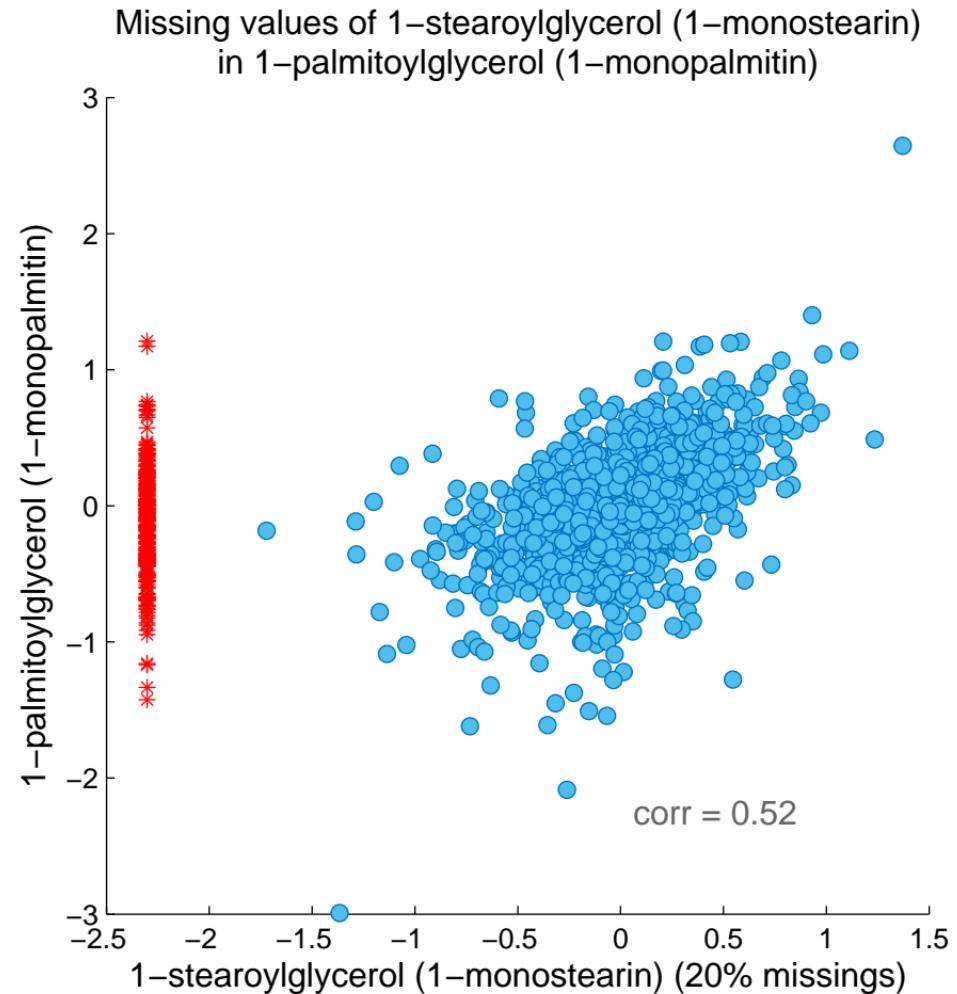




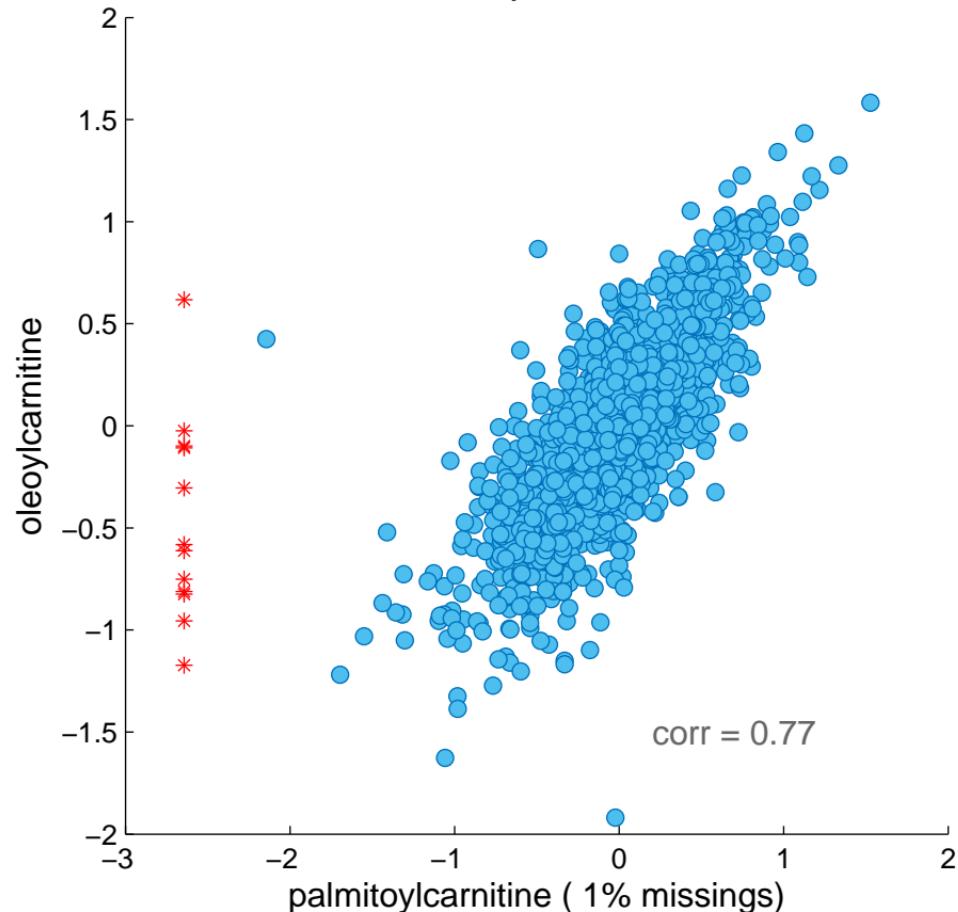




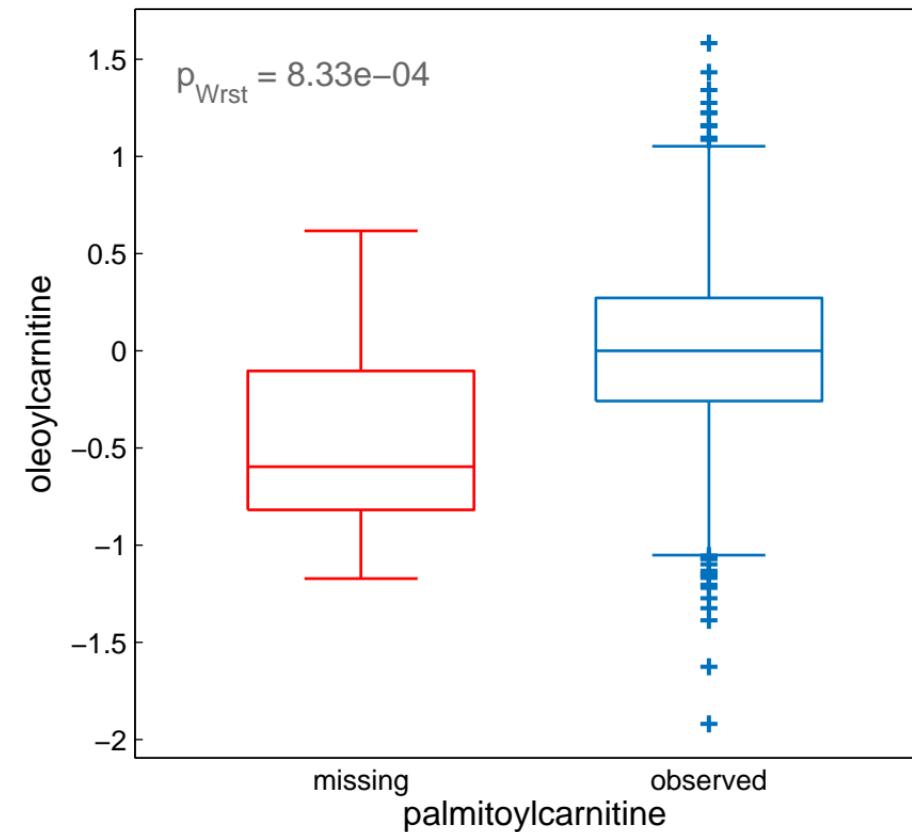


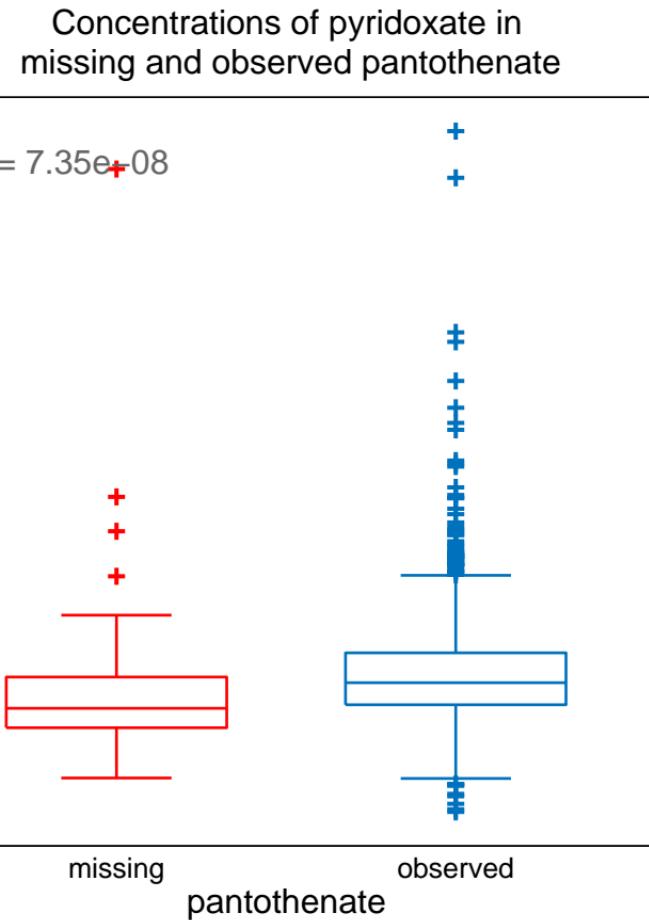
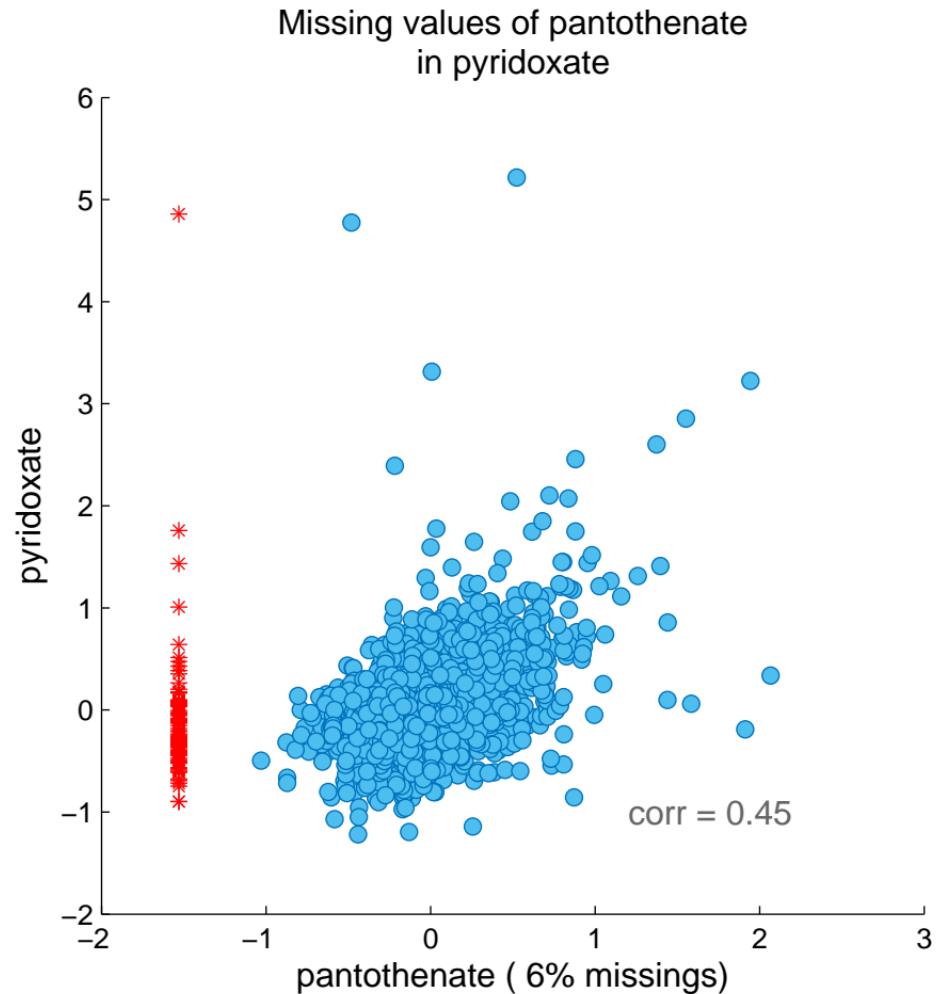


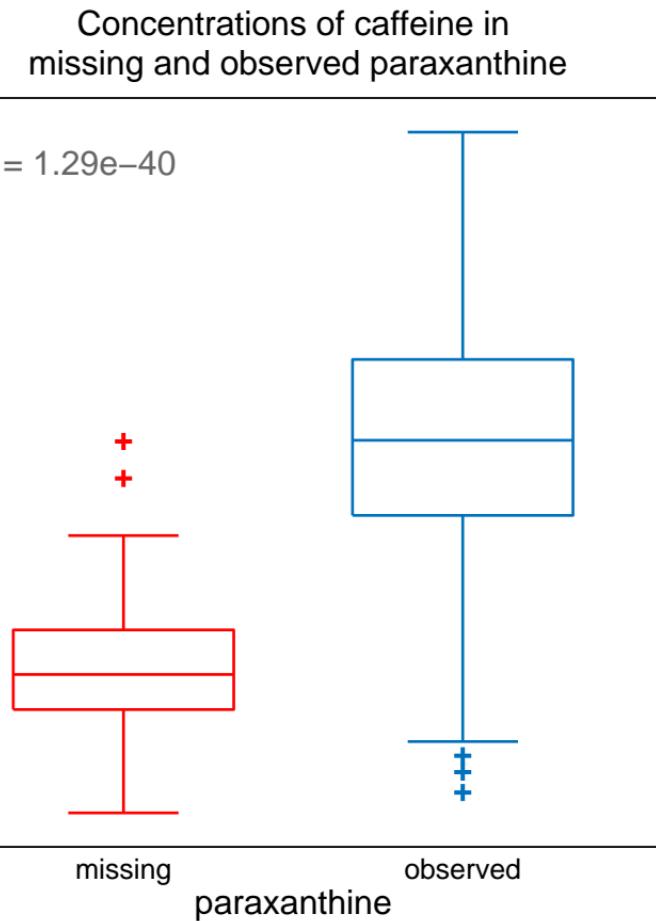
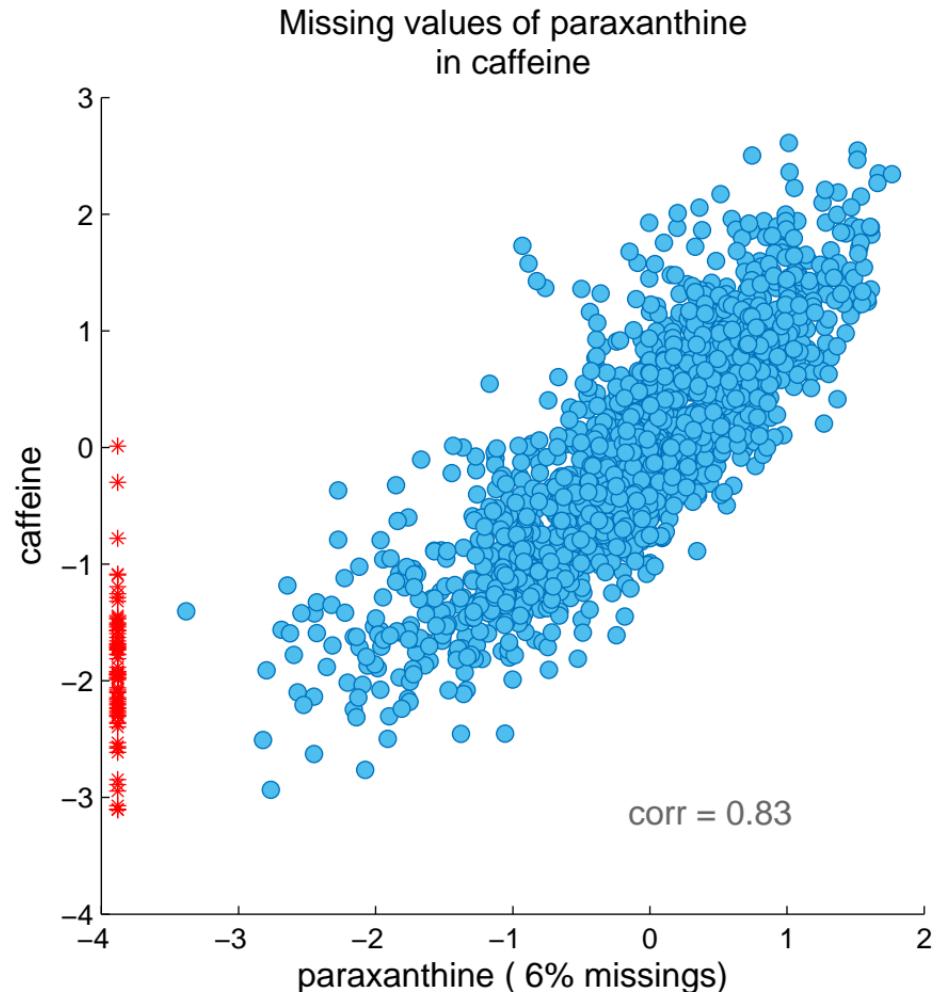
Missing values of palmitoylcarnitine
in oleoylcarnitine

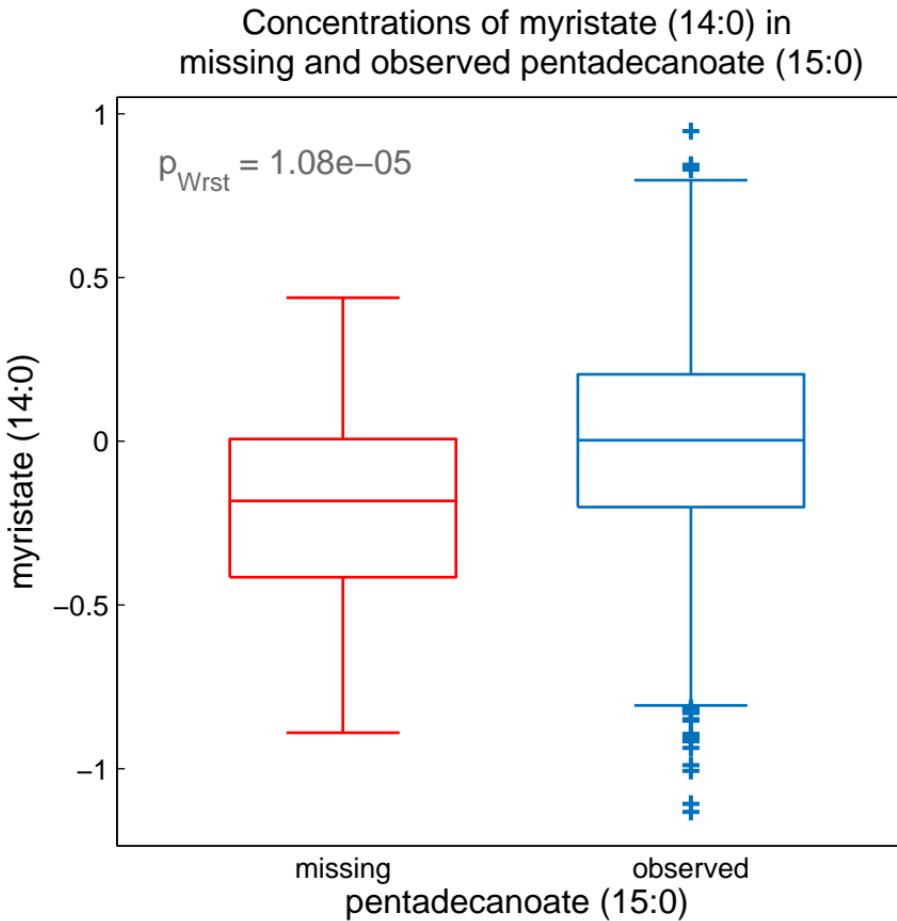
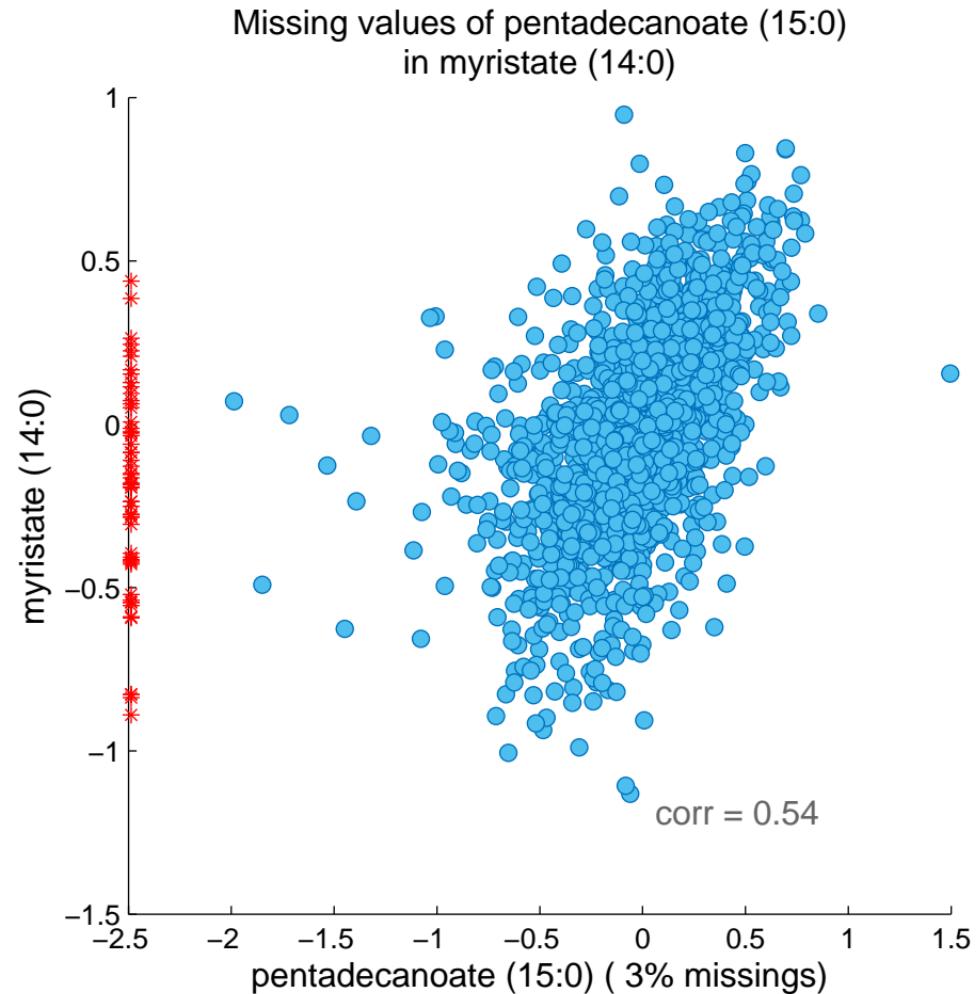


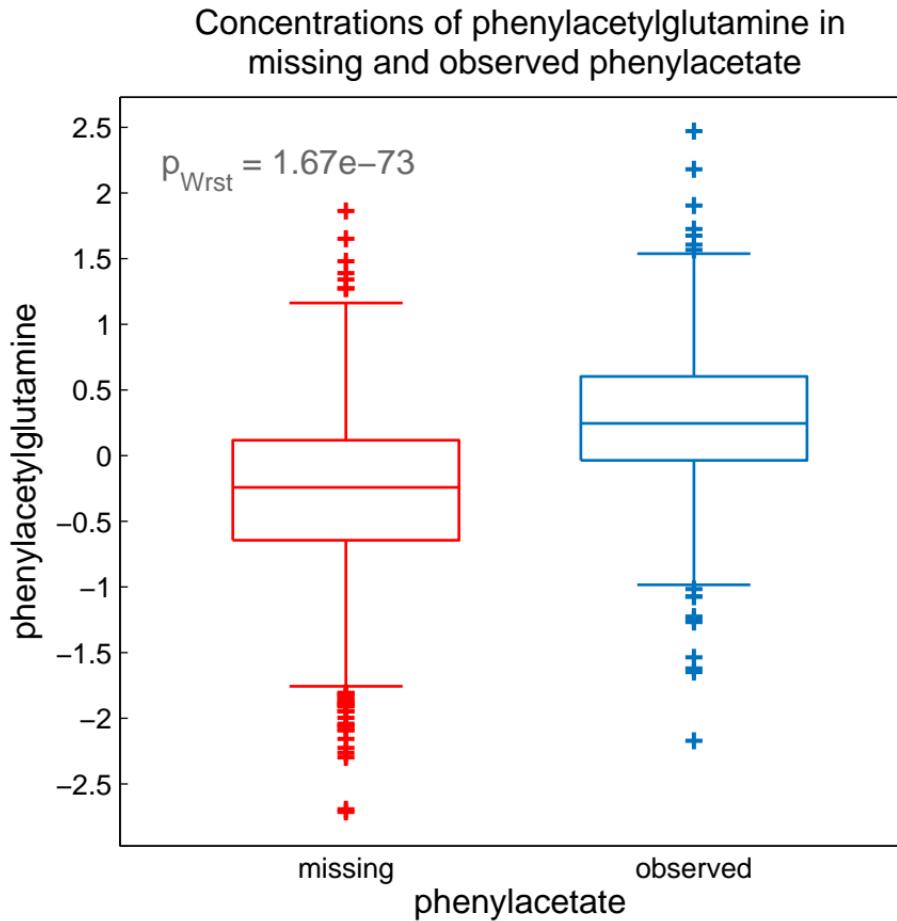
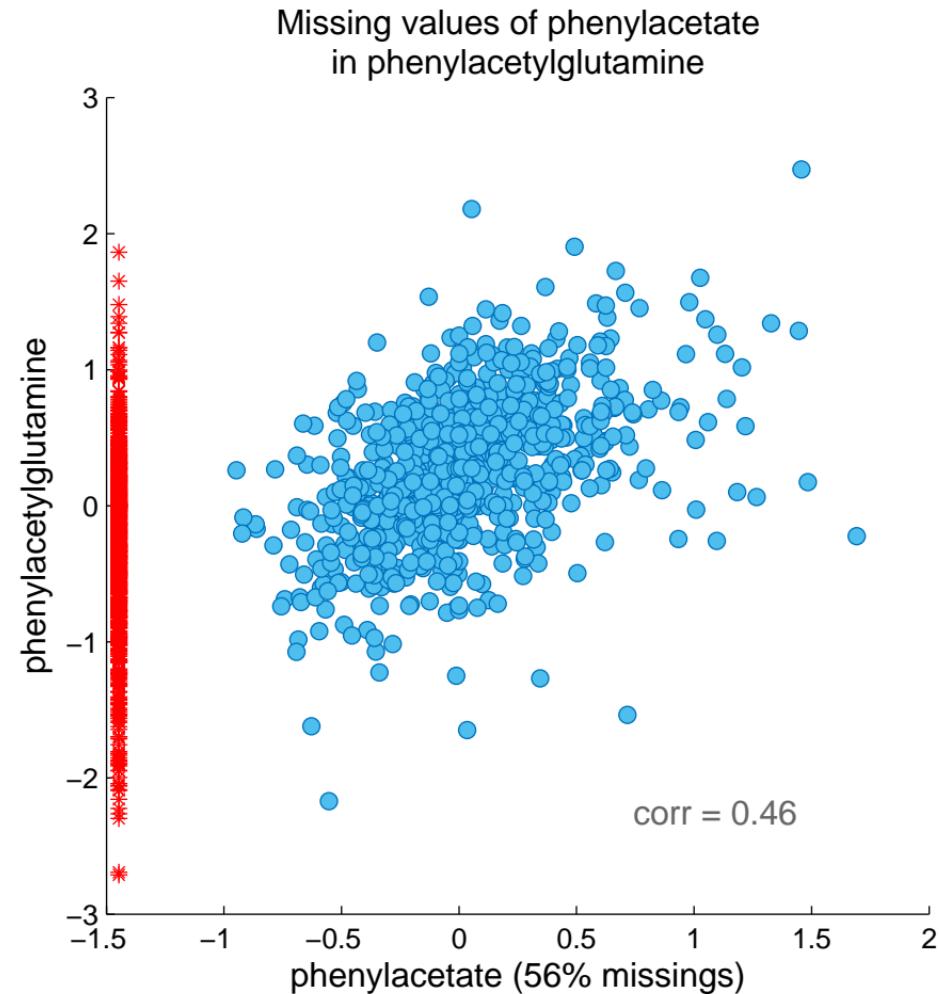
Concentrations of oleoylcarnitine in
missing and observed palmitoylcarnitine

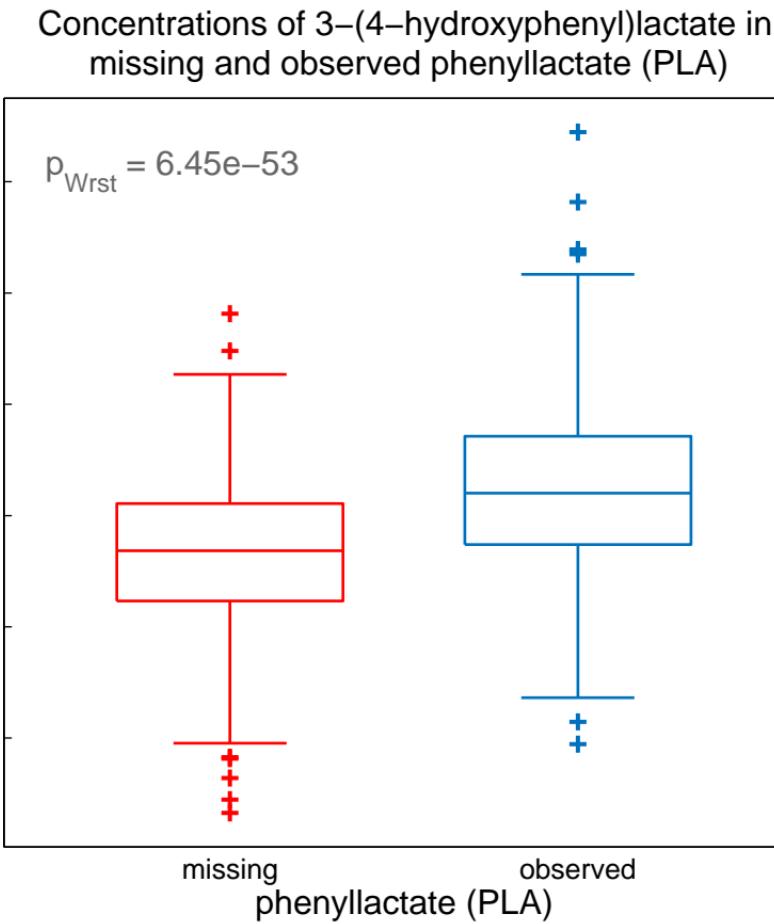
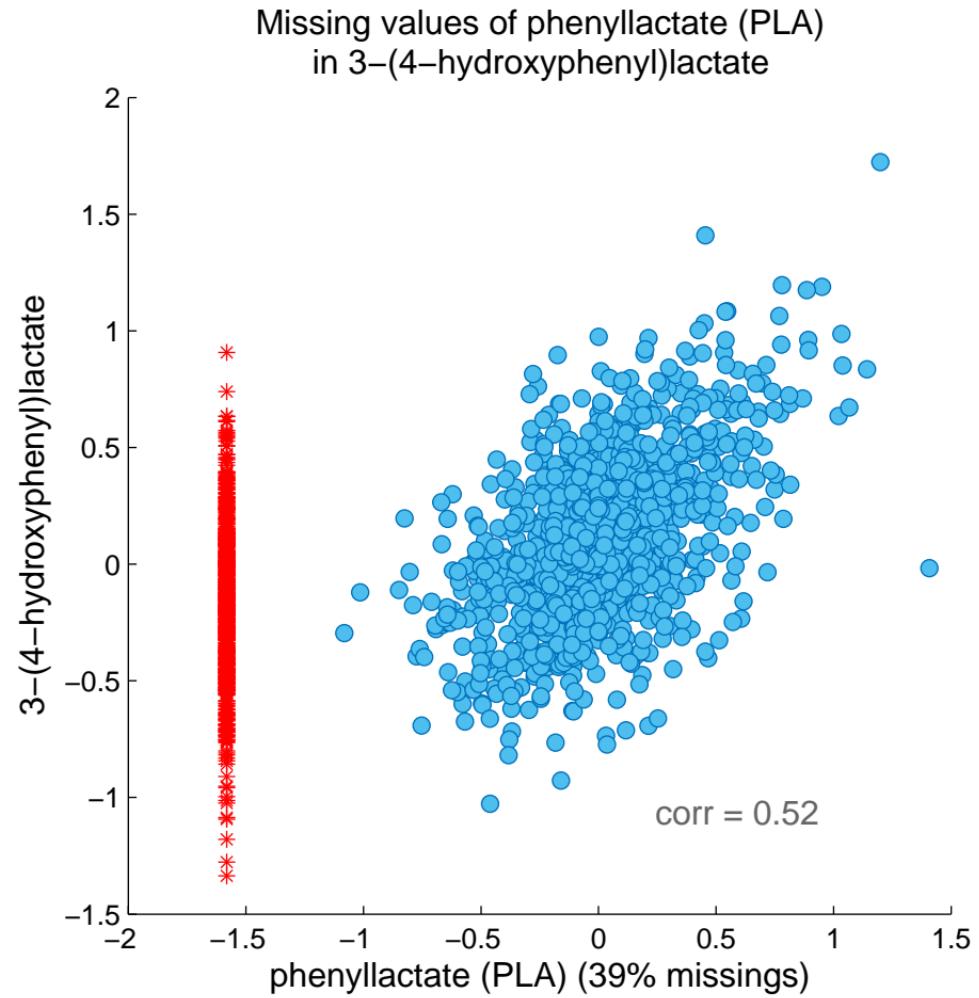


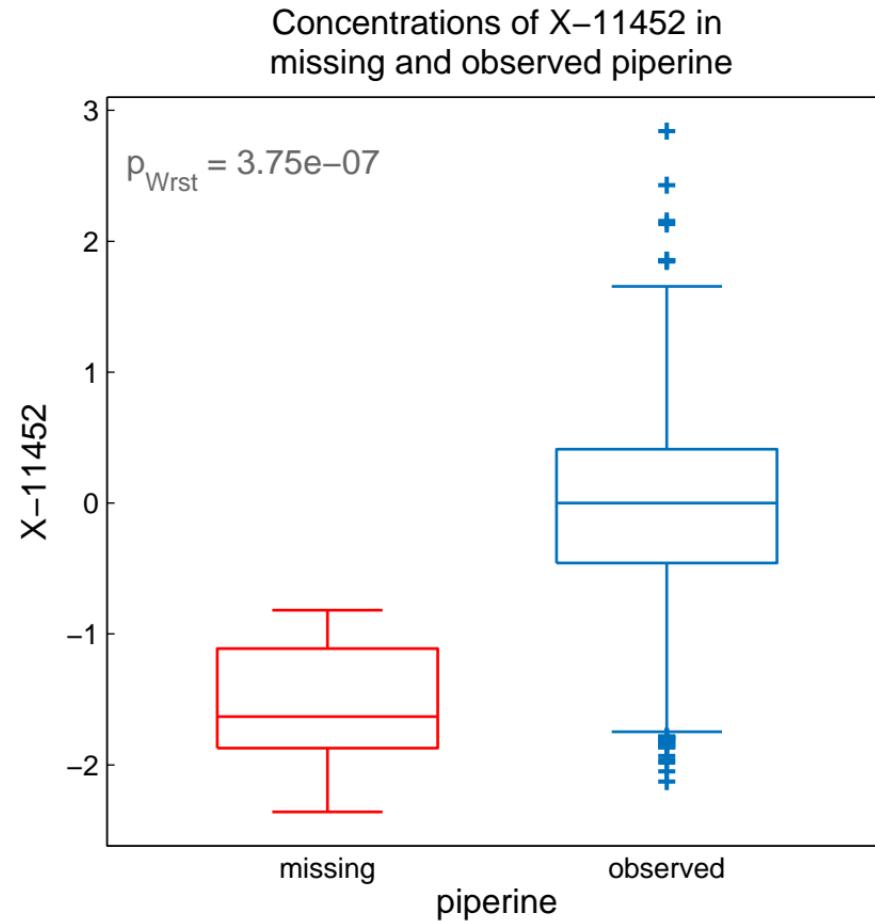
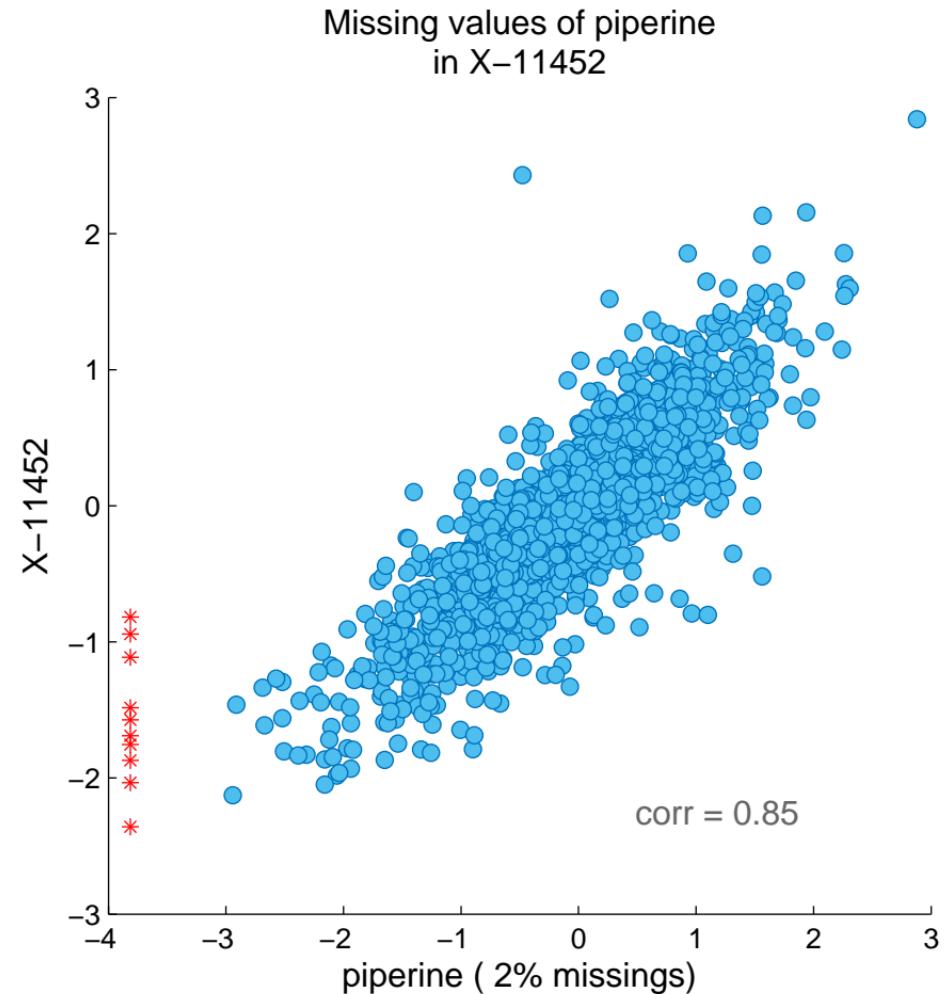


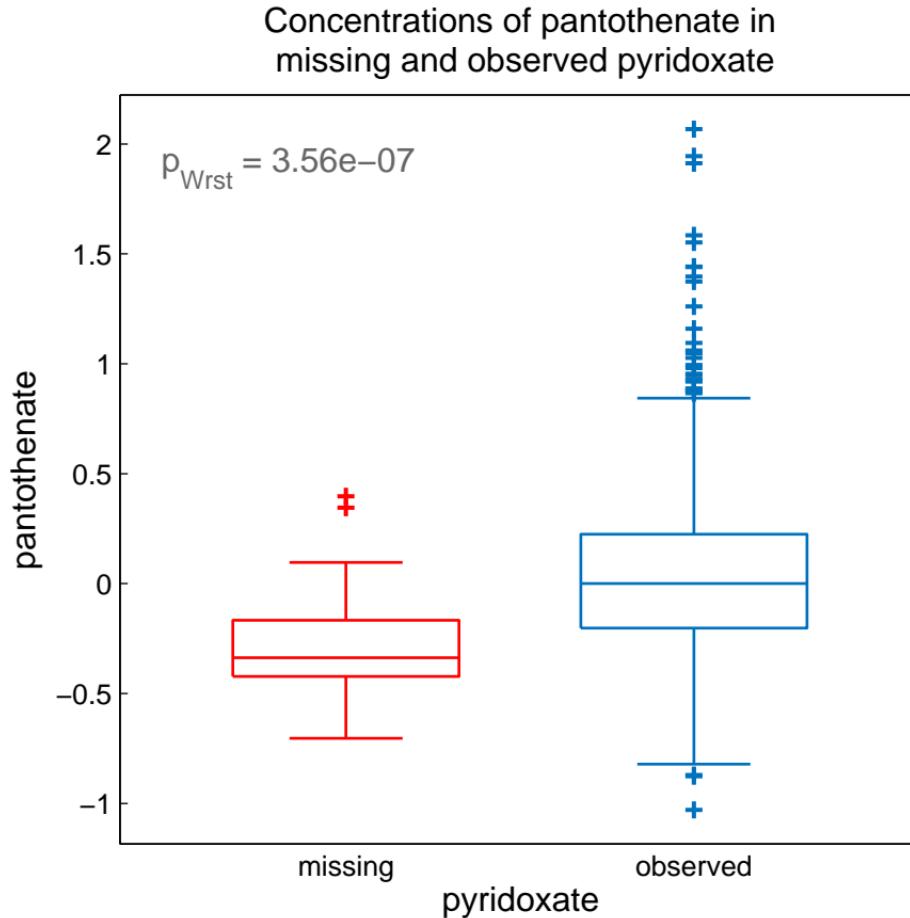
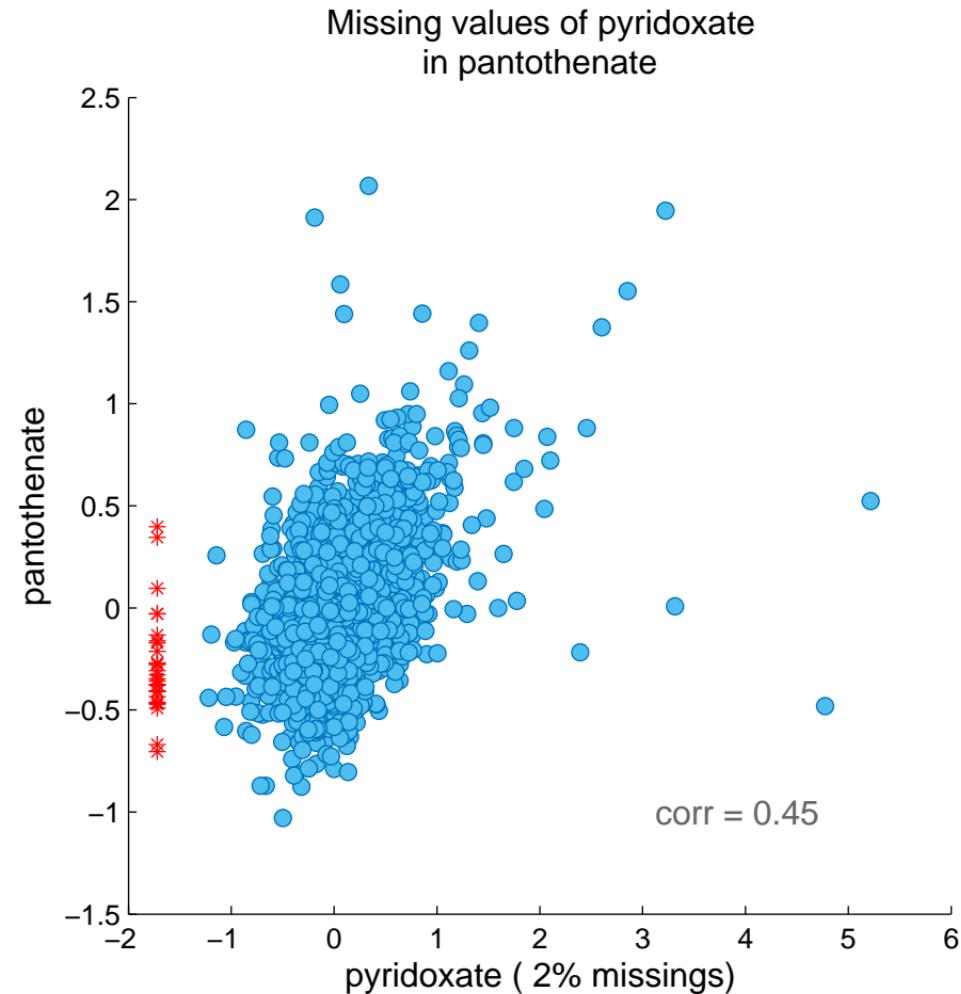


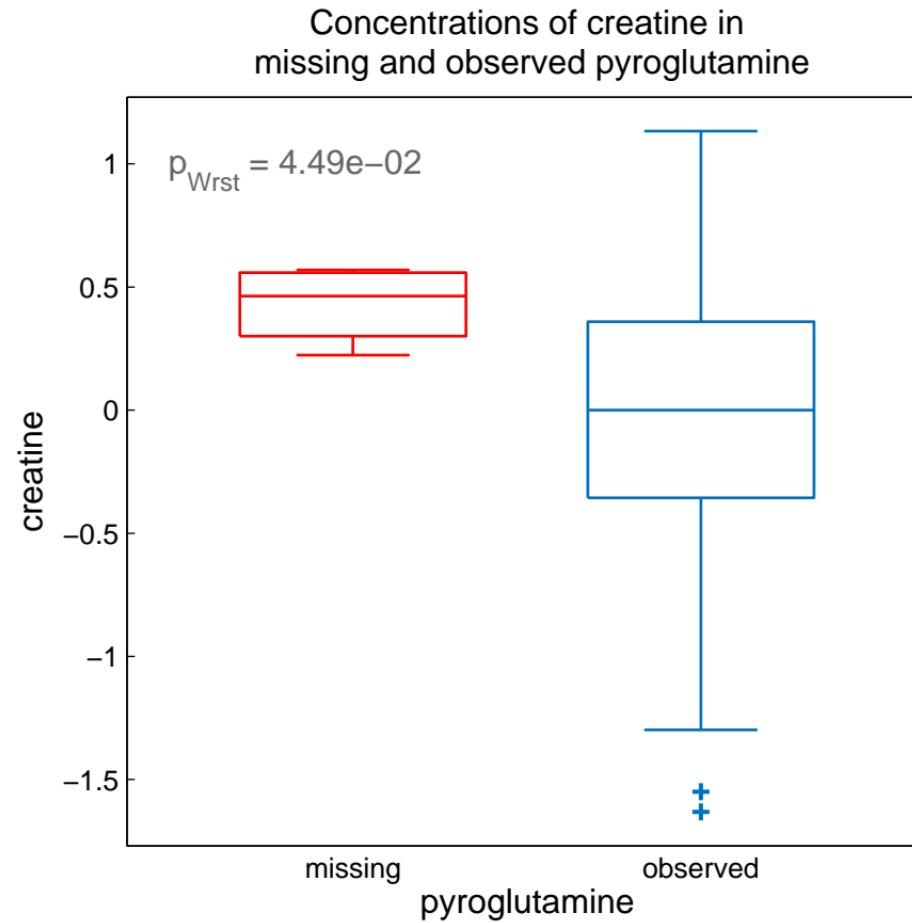
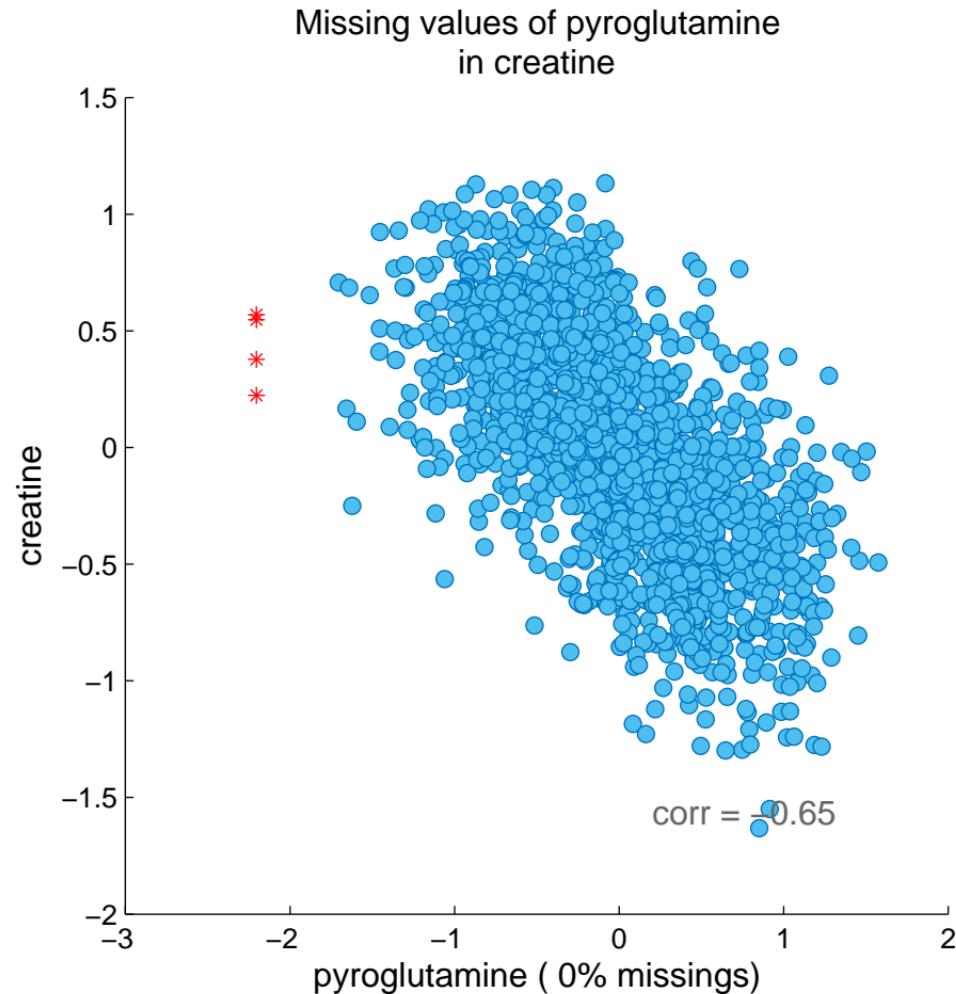


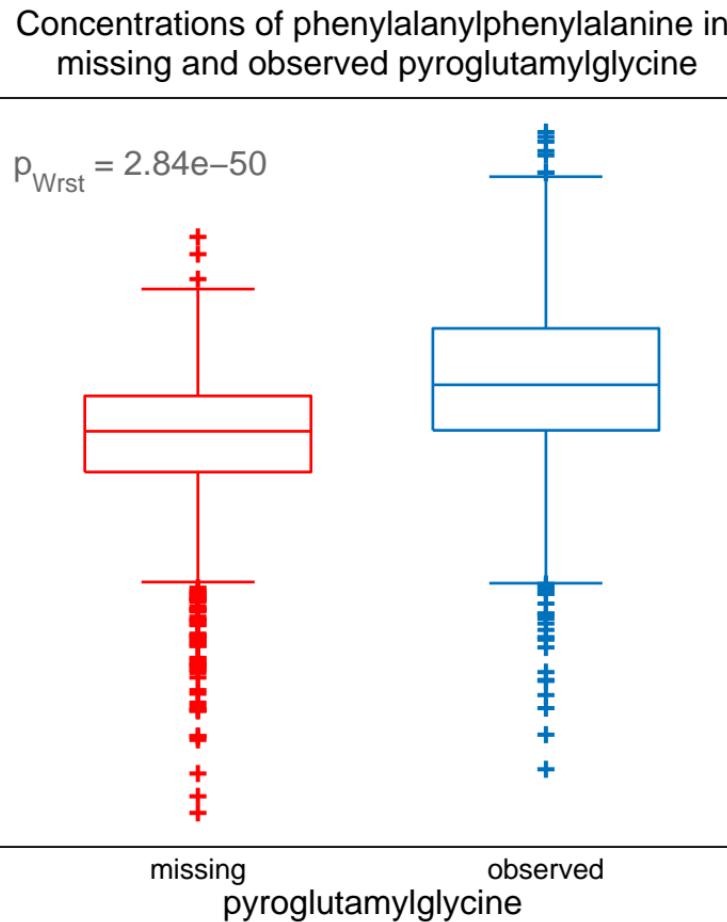
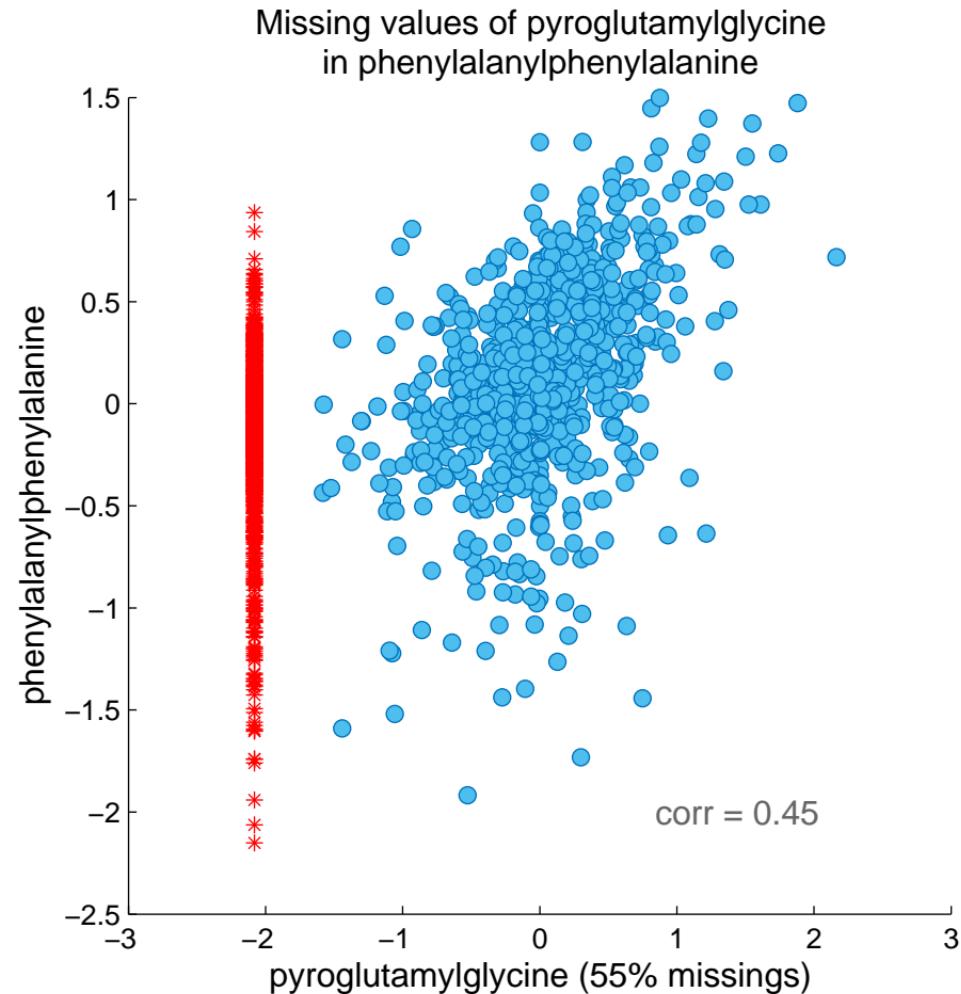


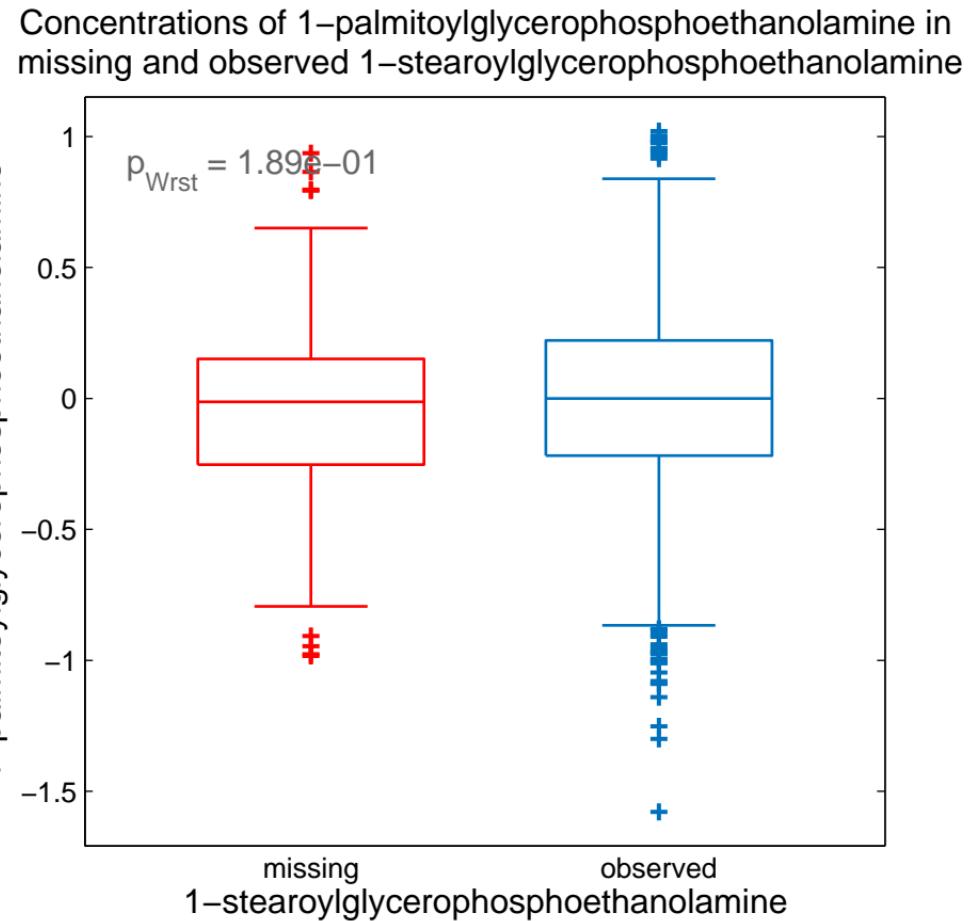
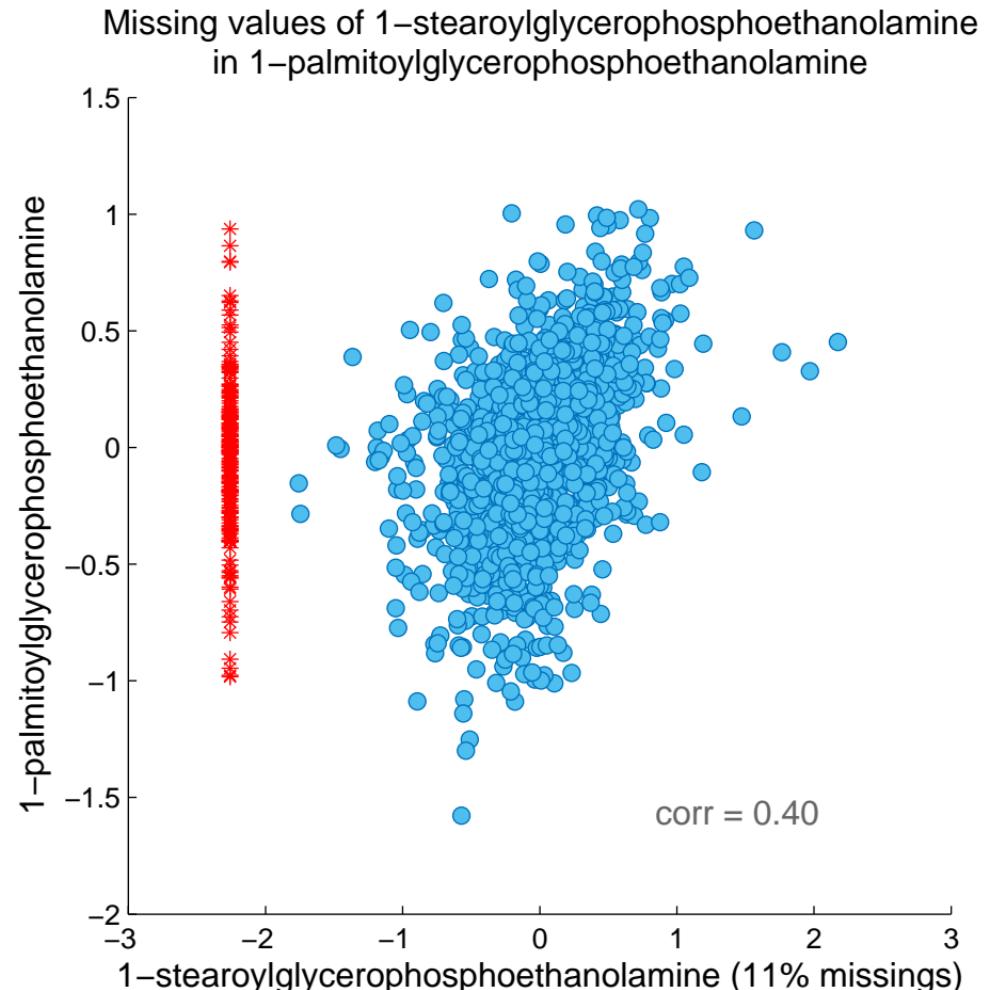


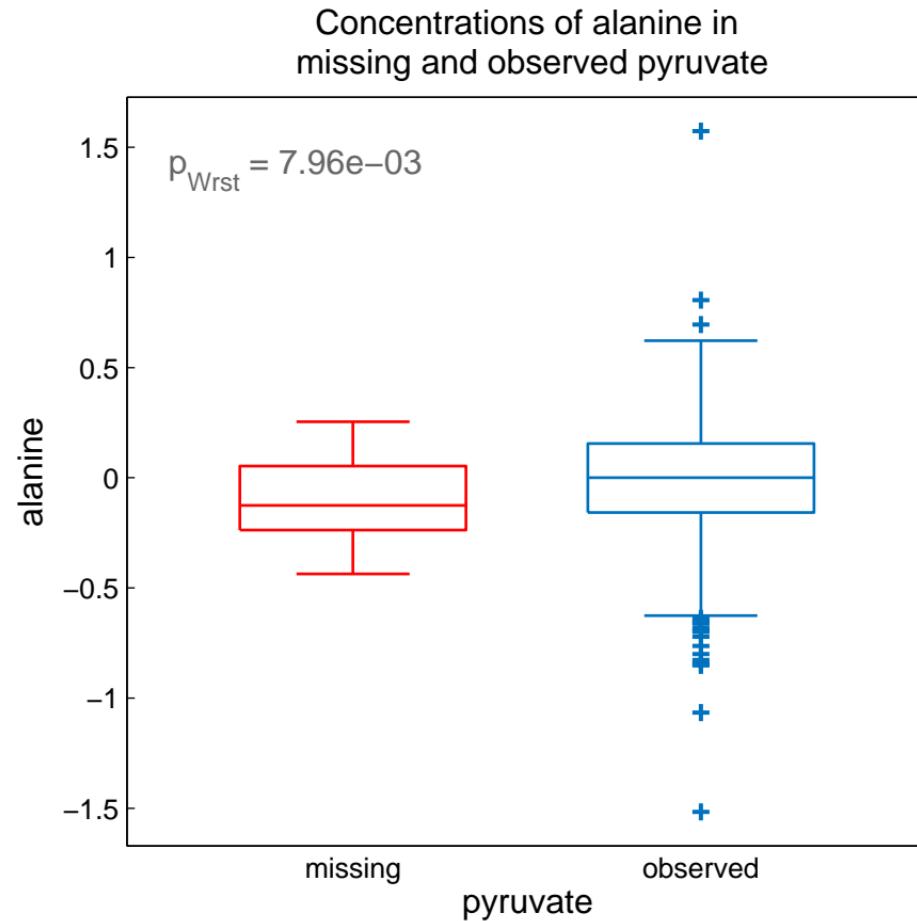
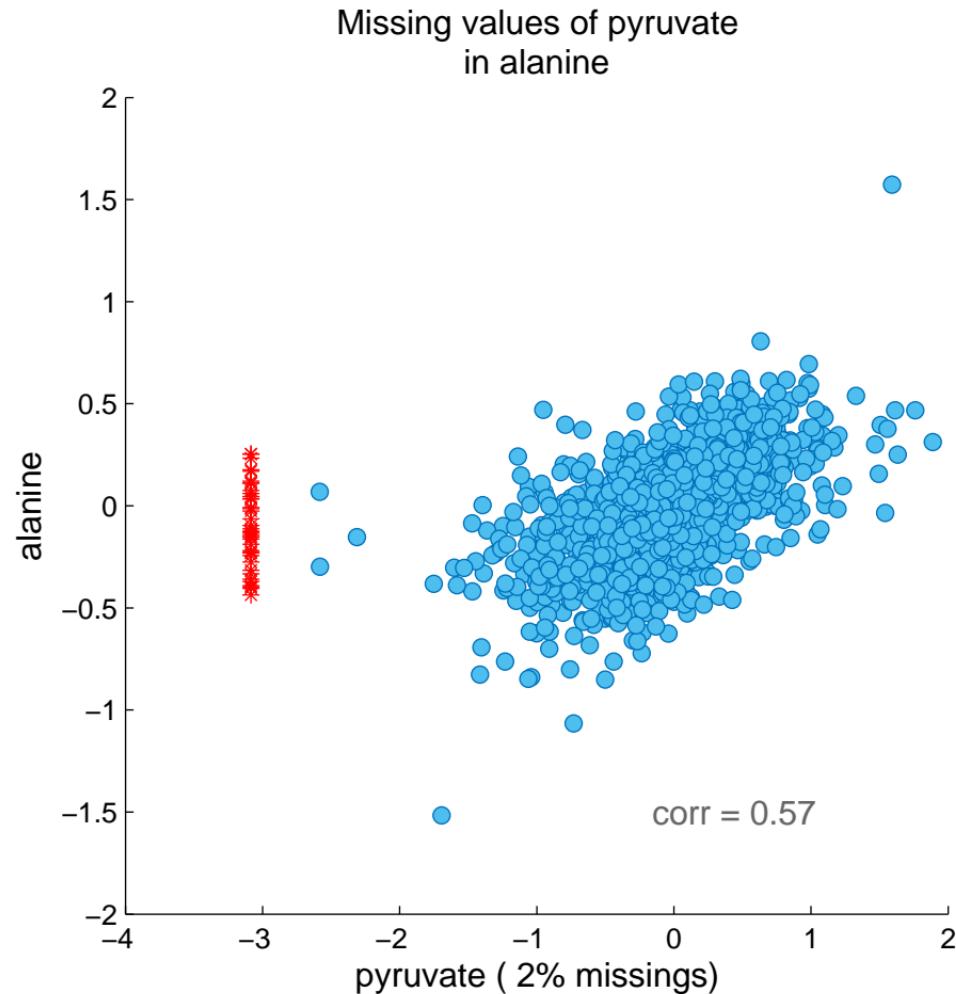


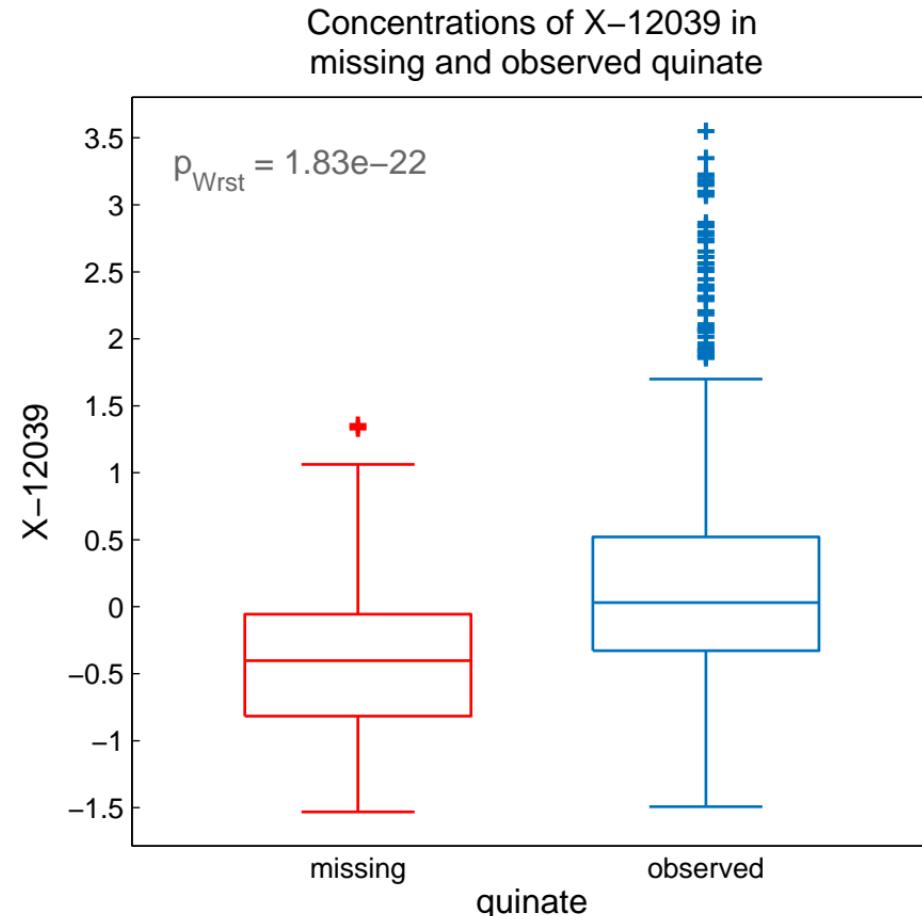
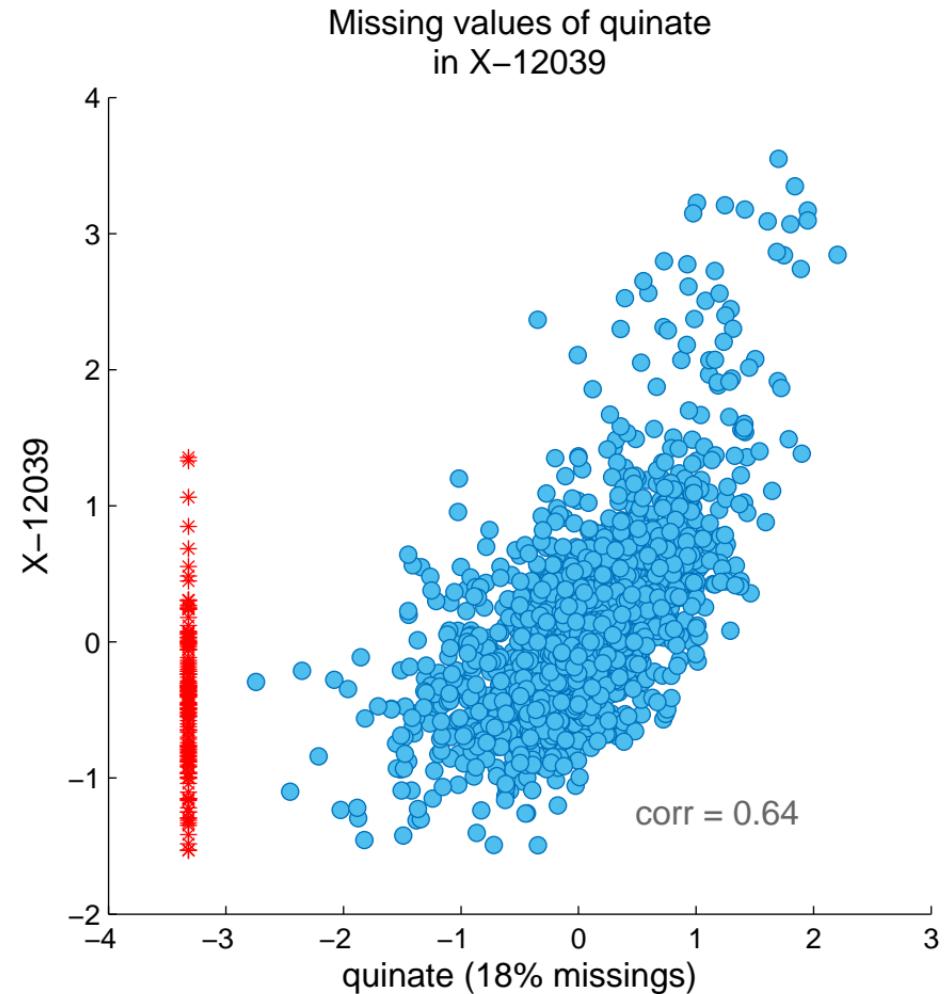


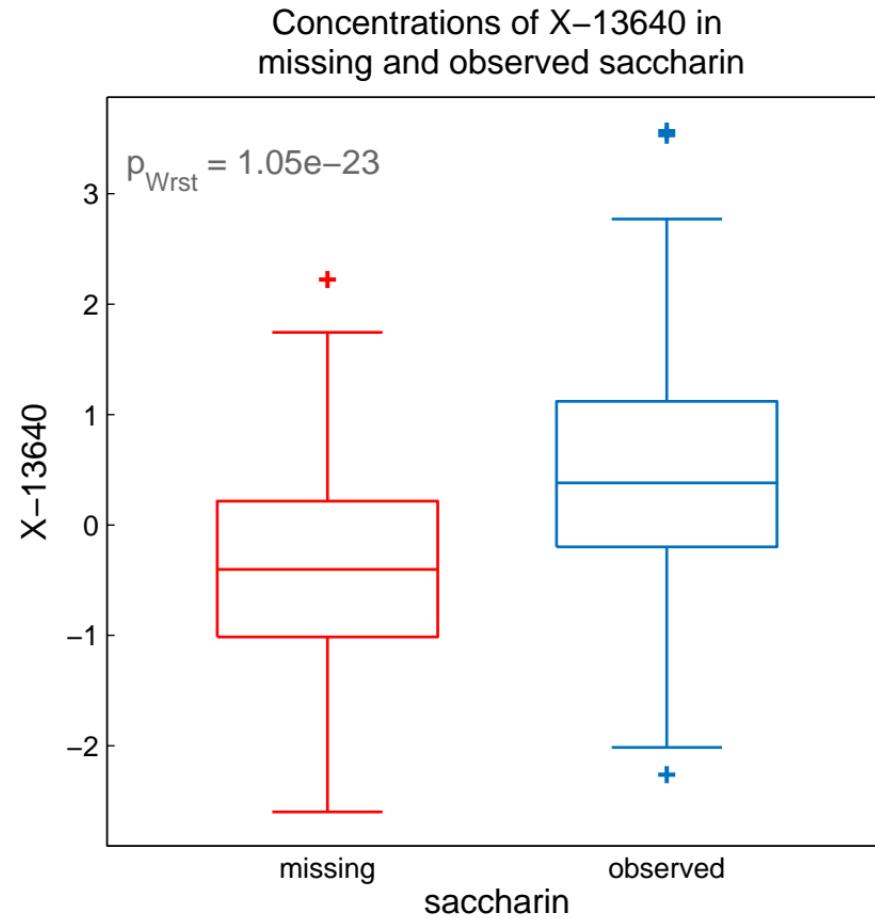
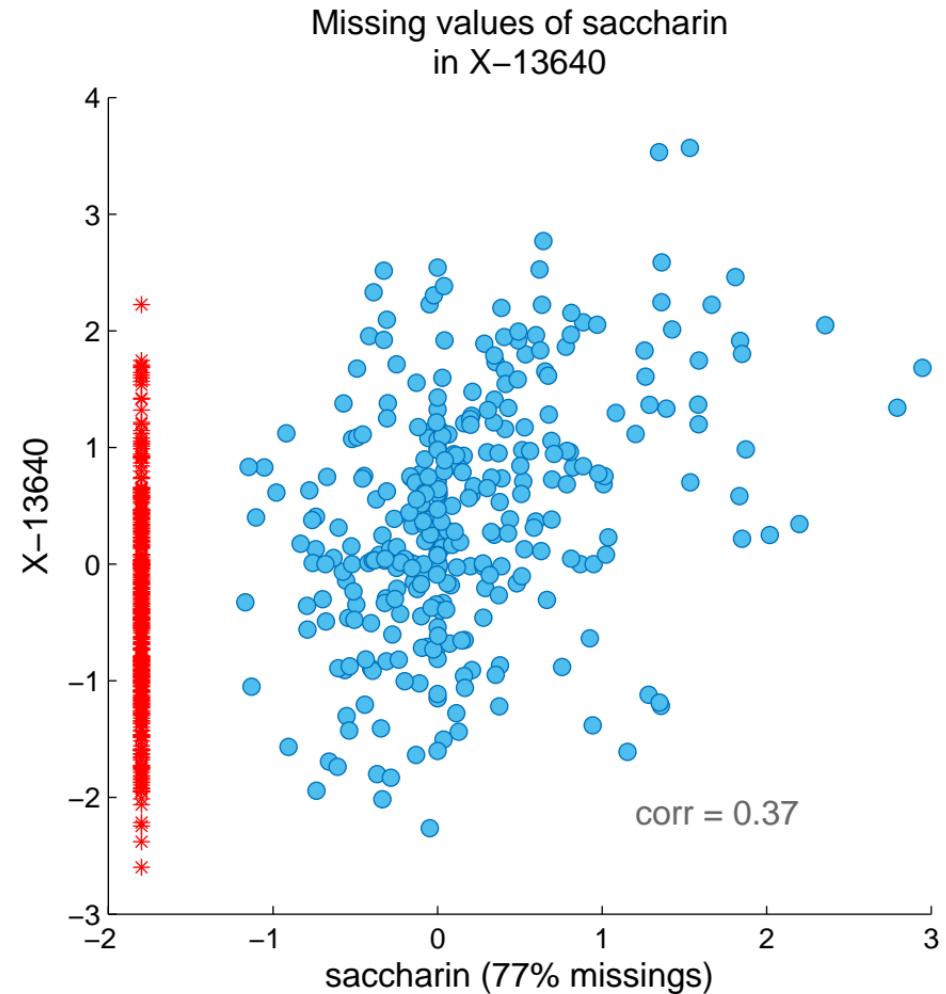


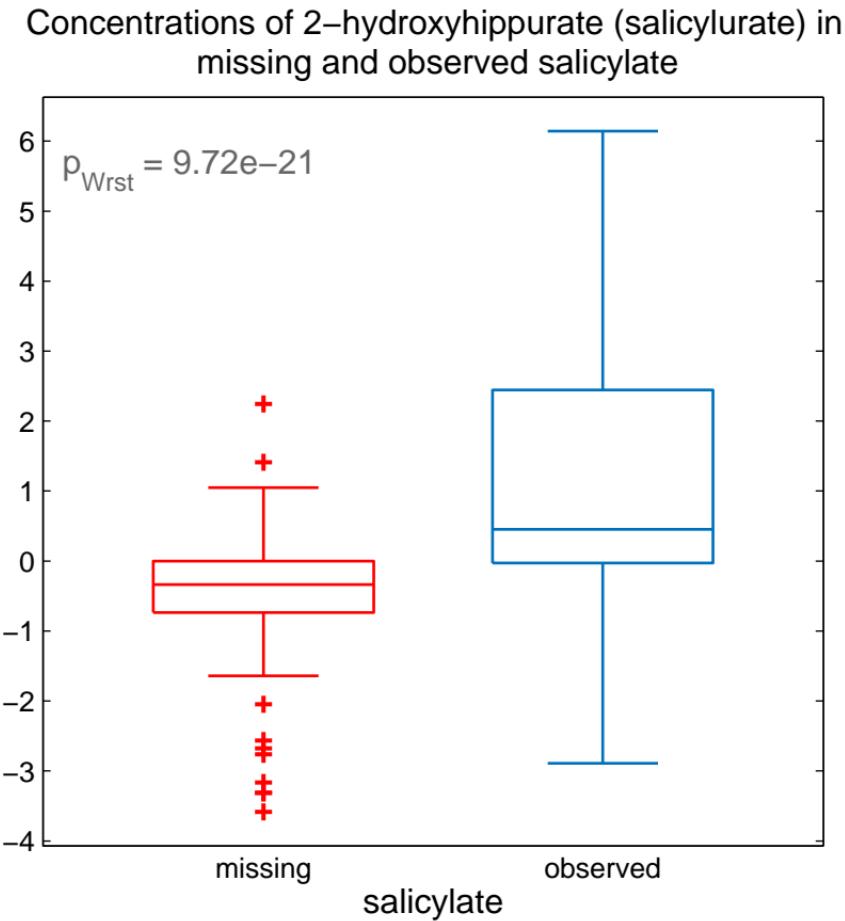
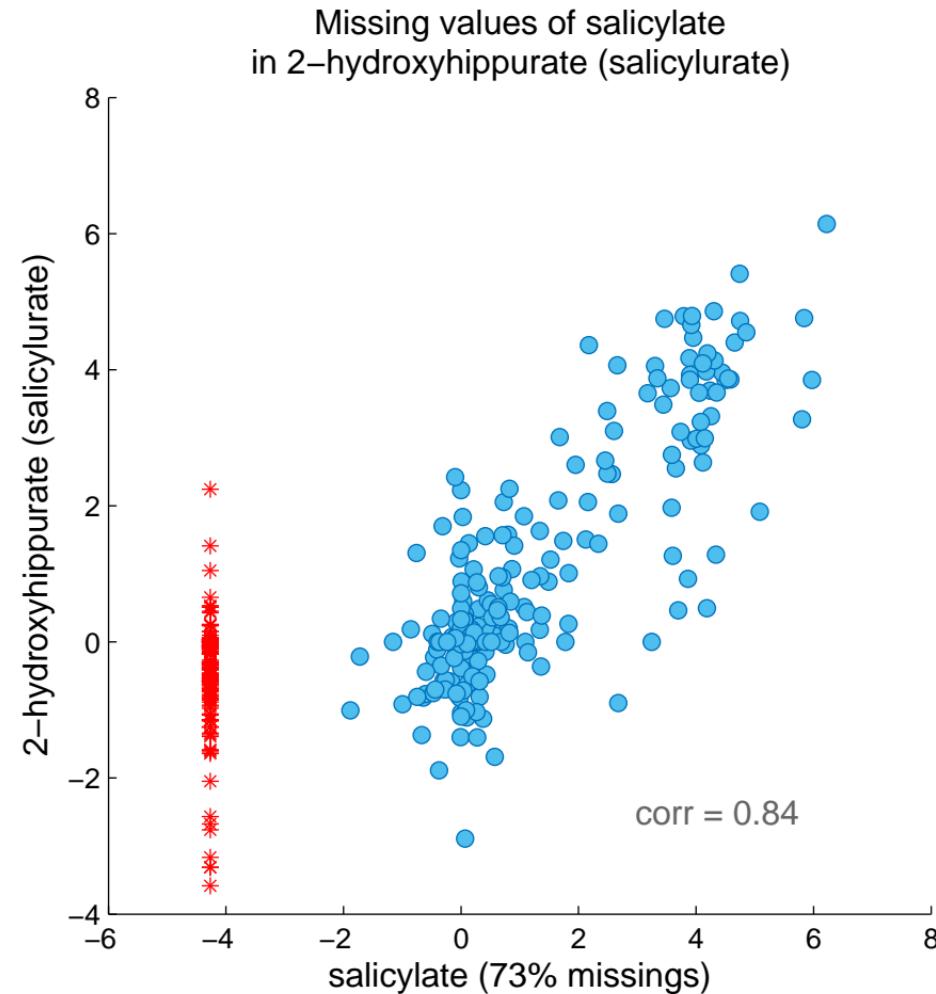


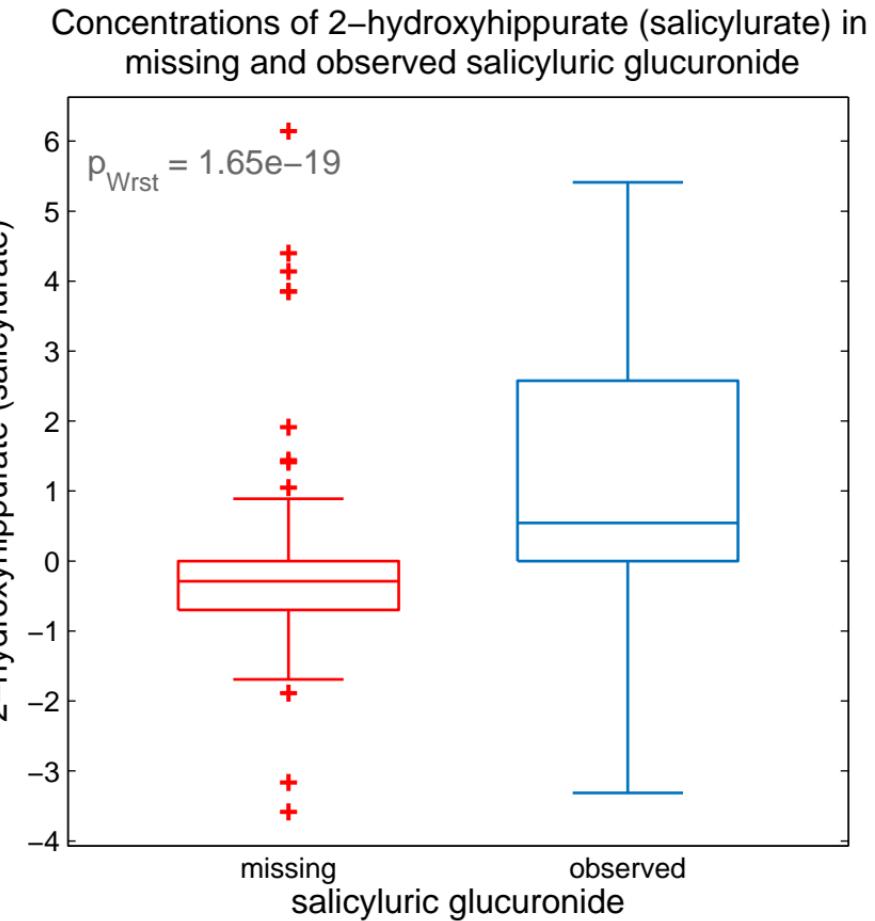
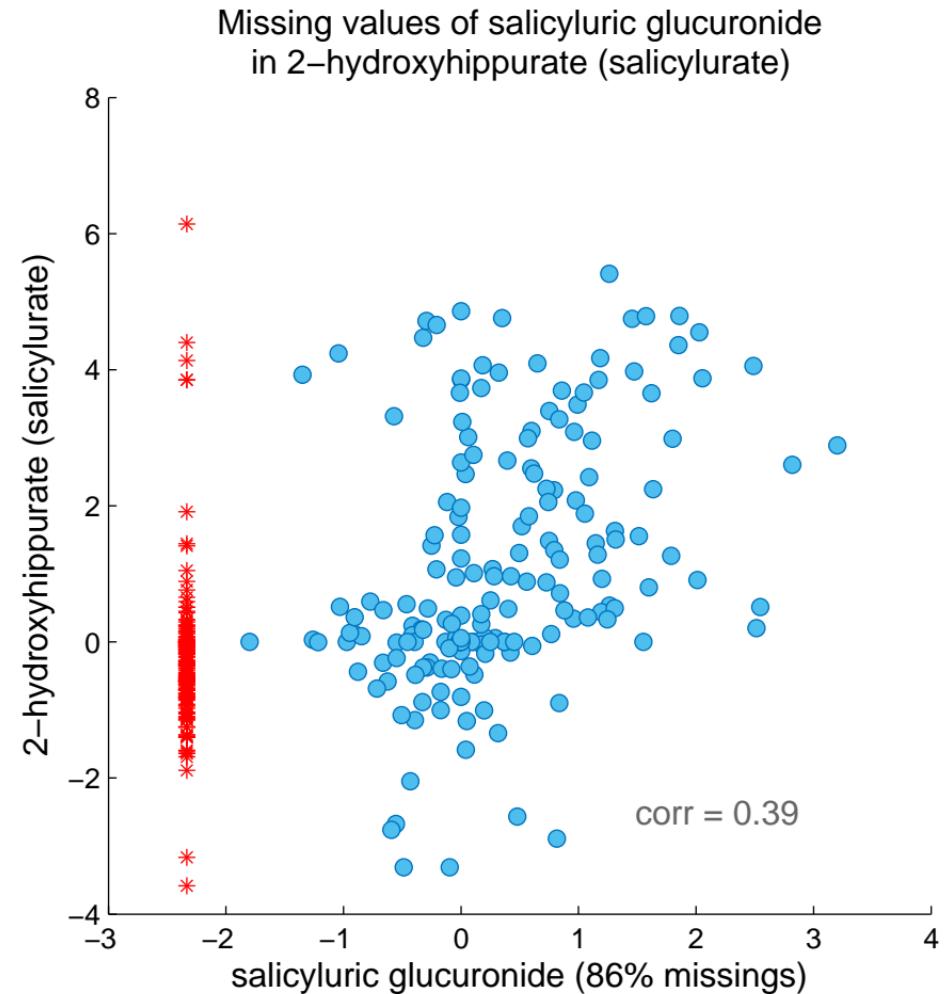


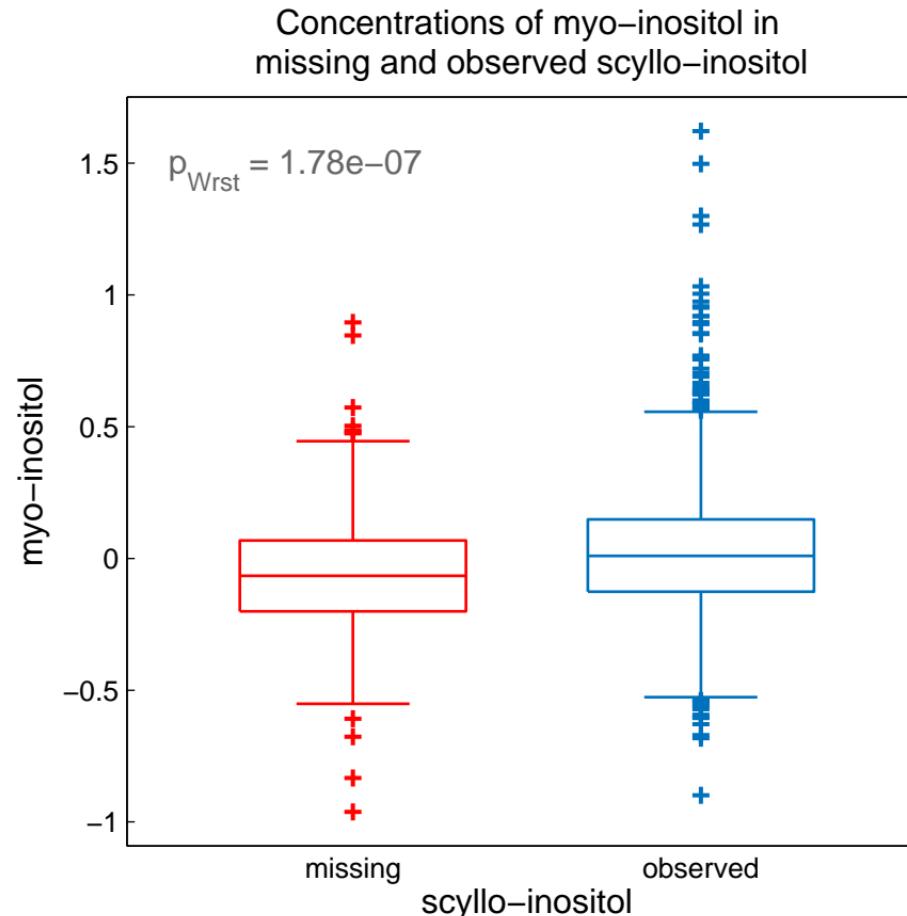
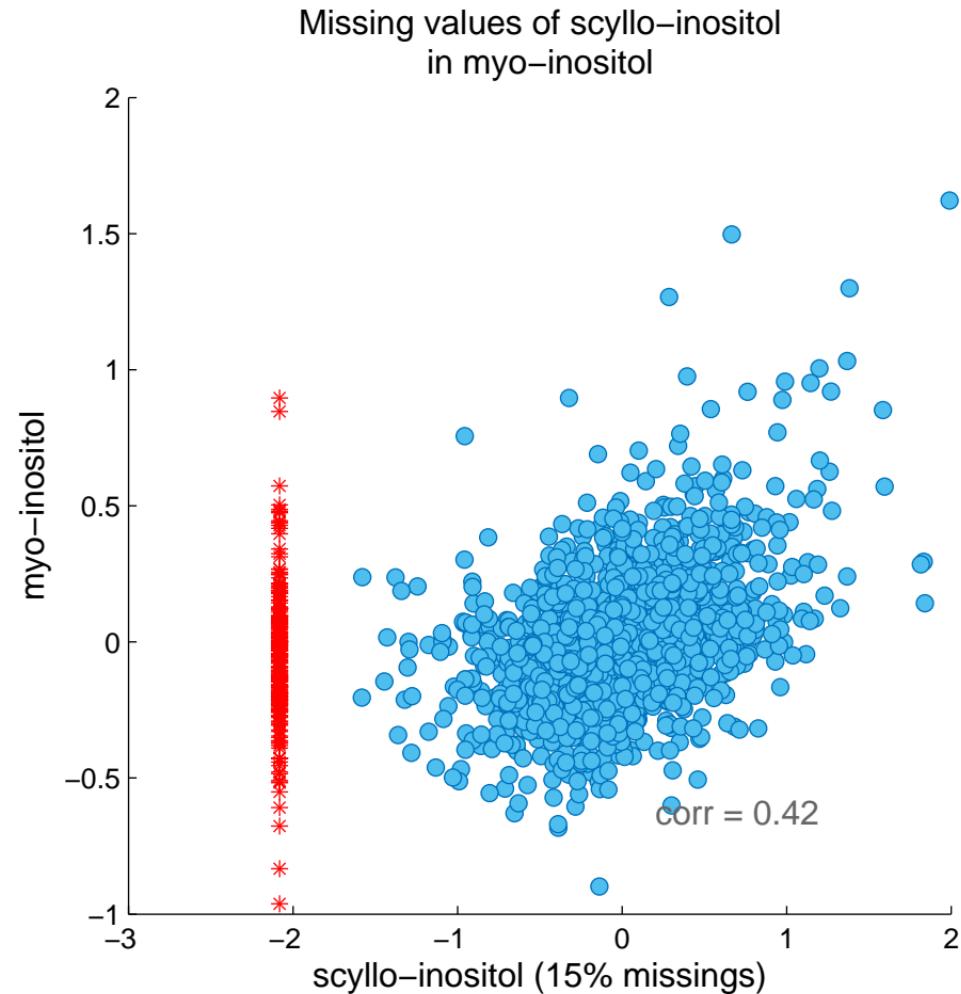


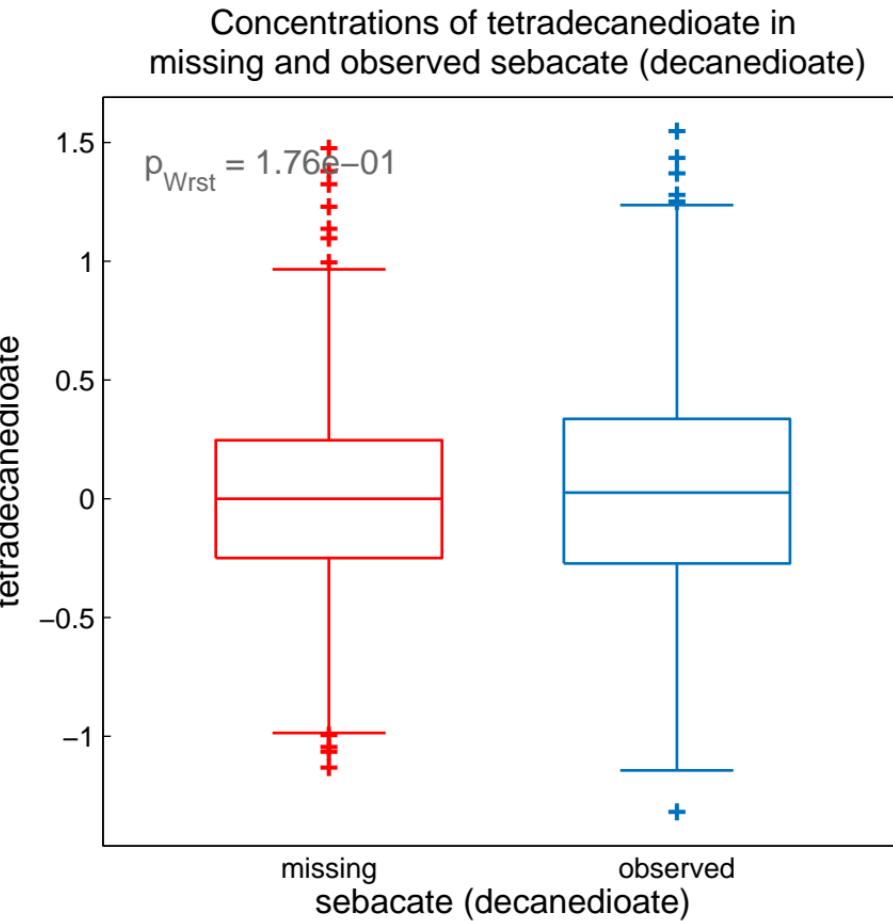
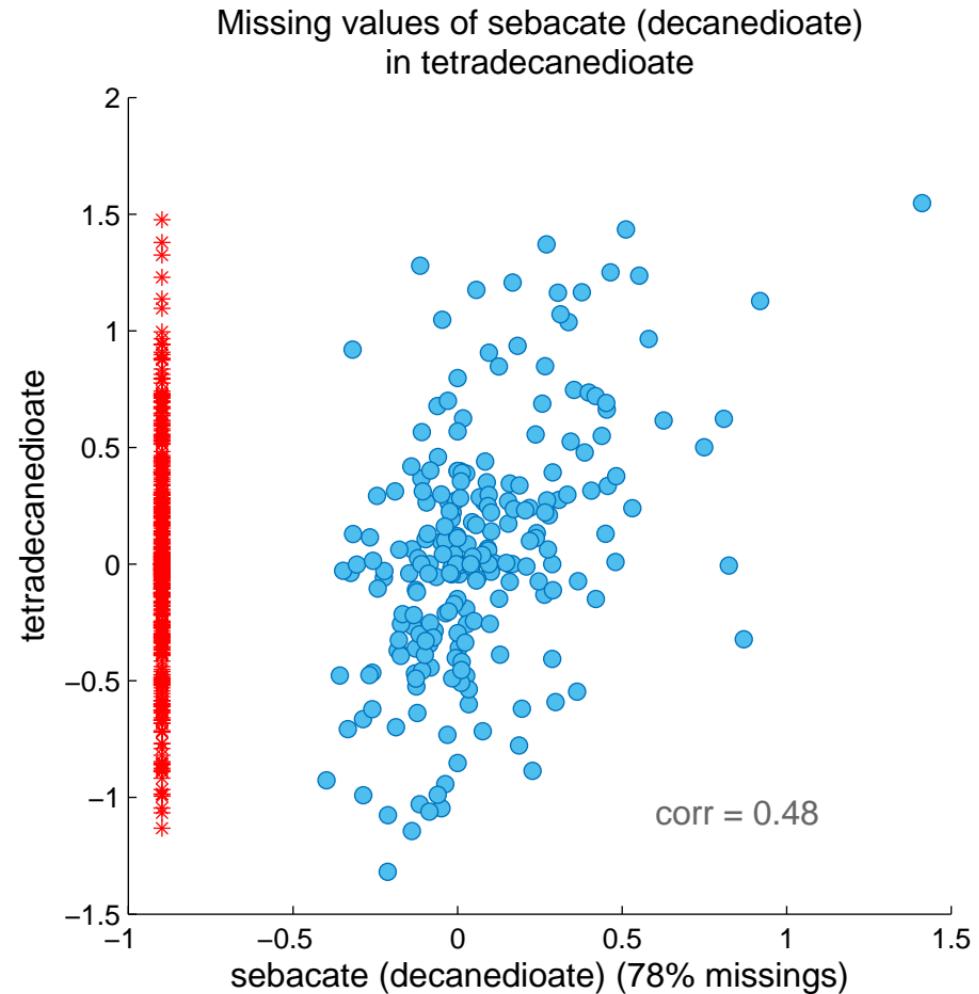


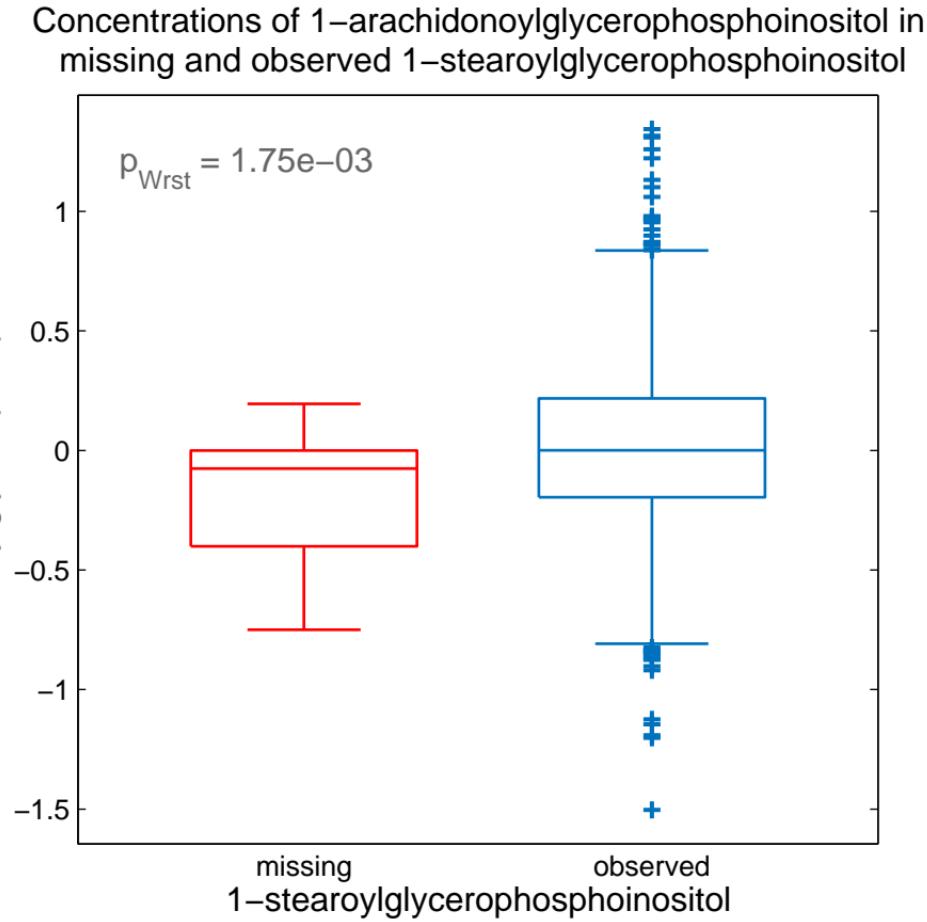
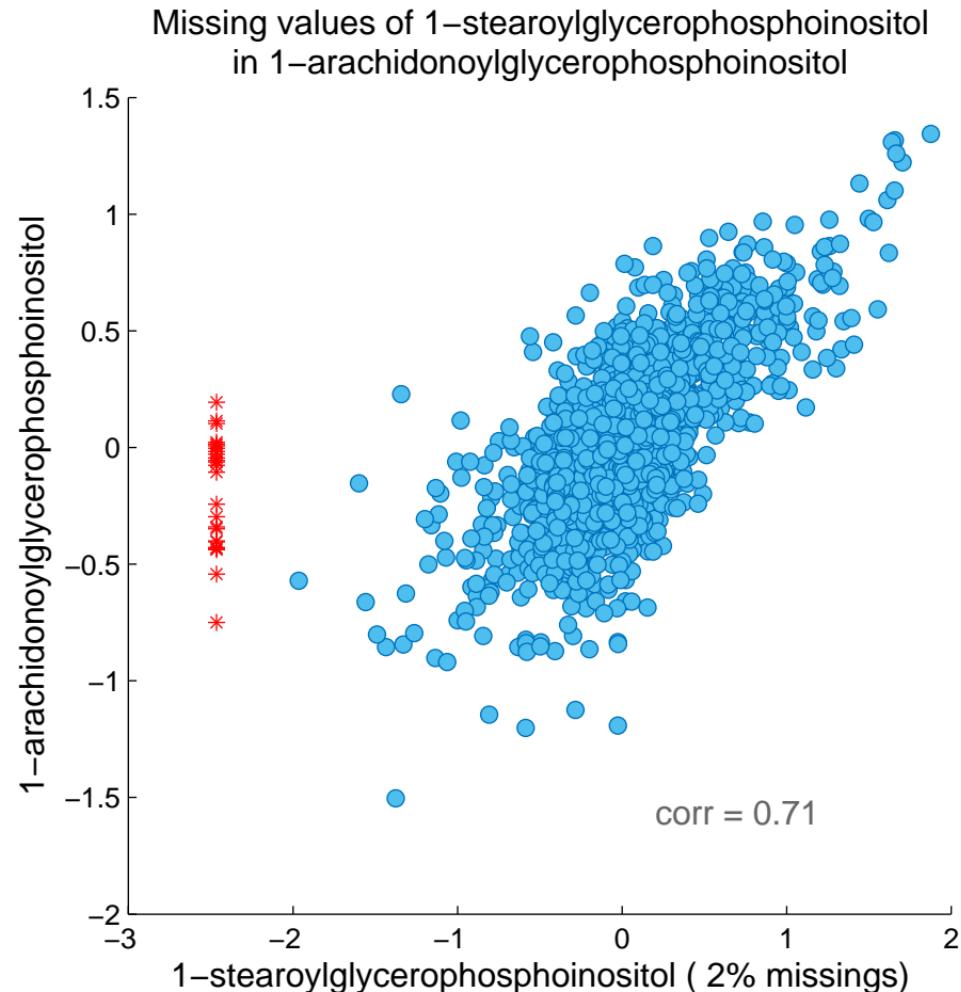


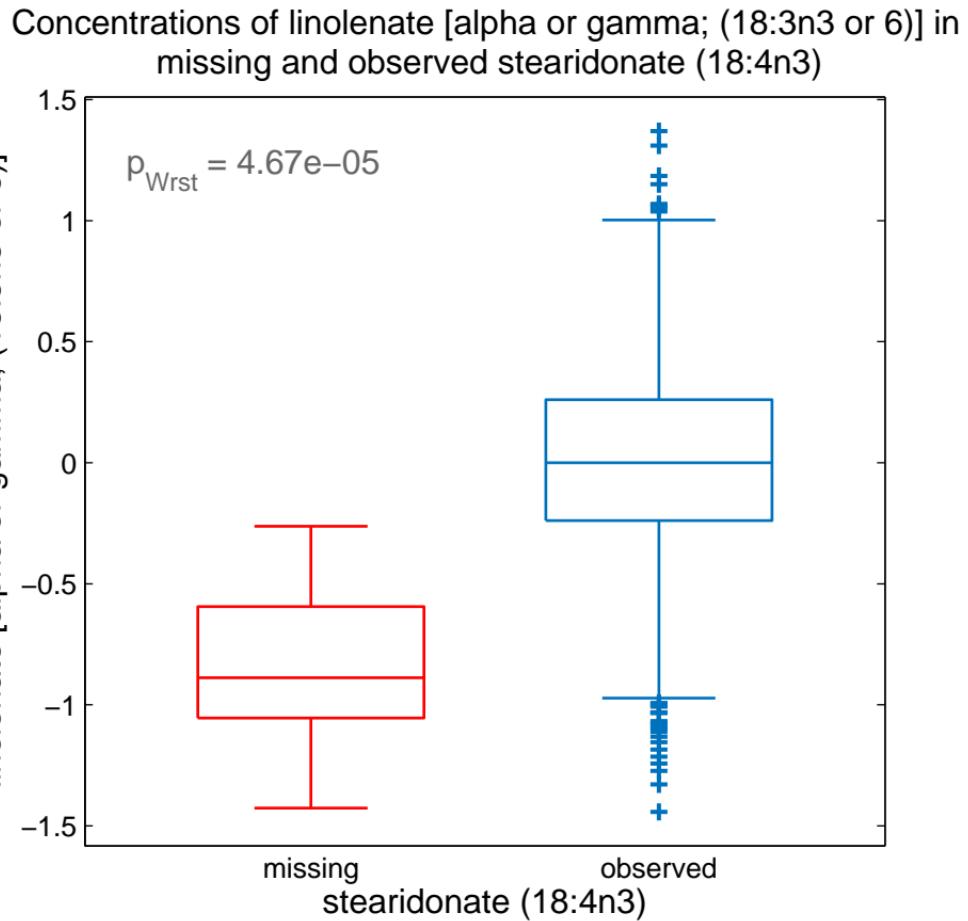
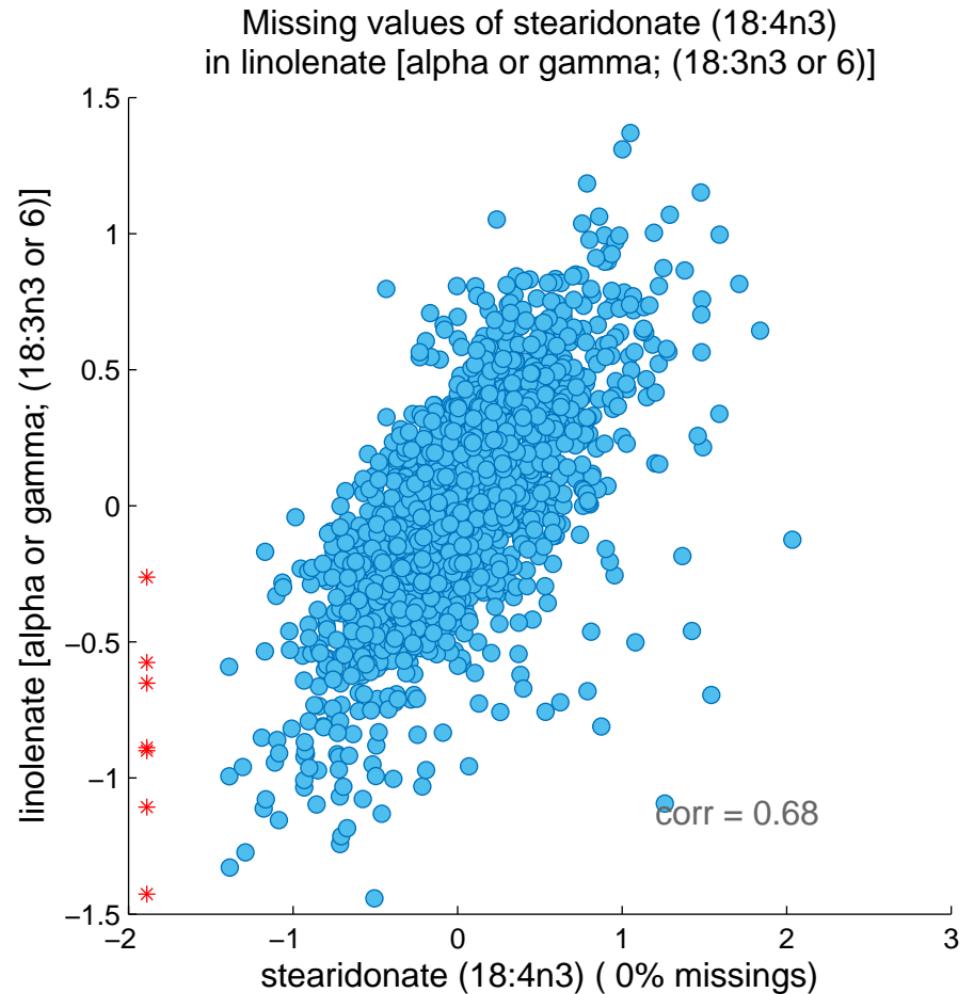




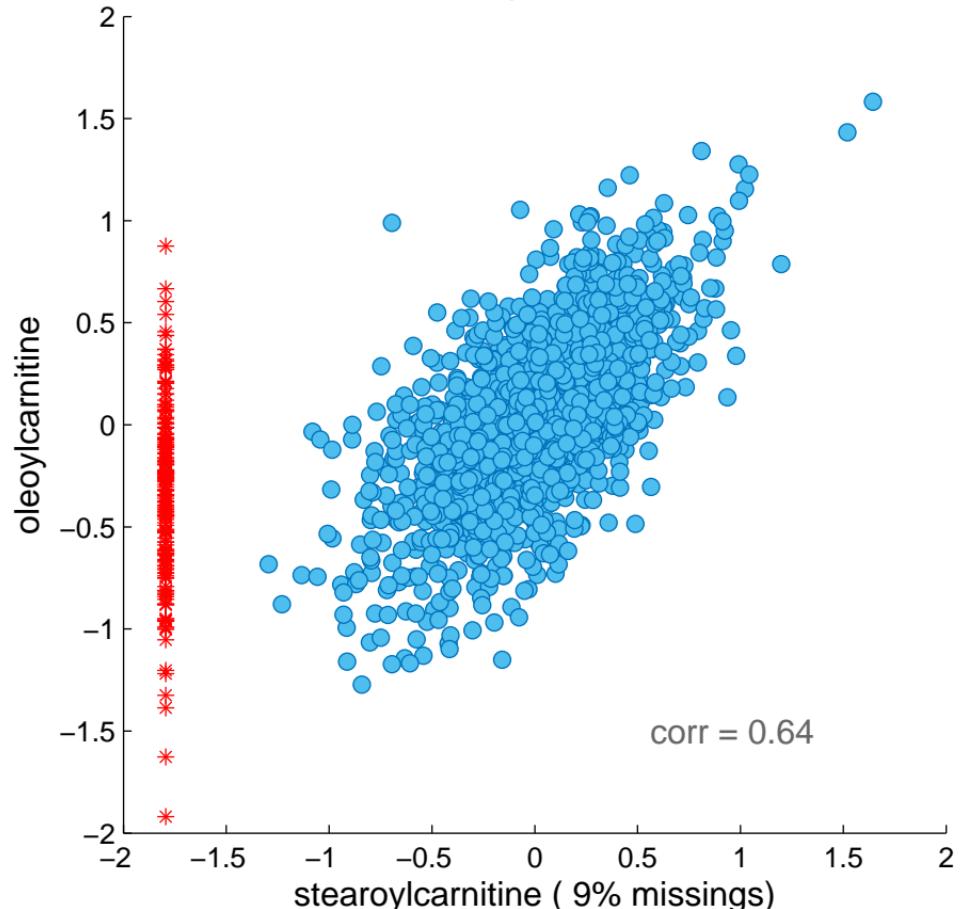




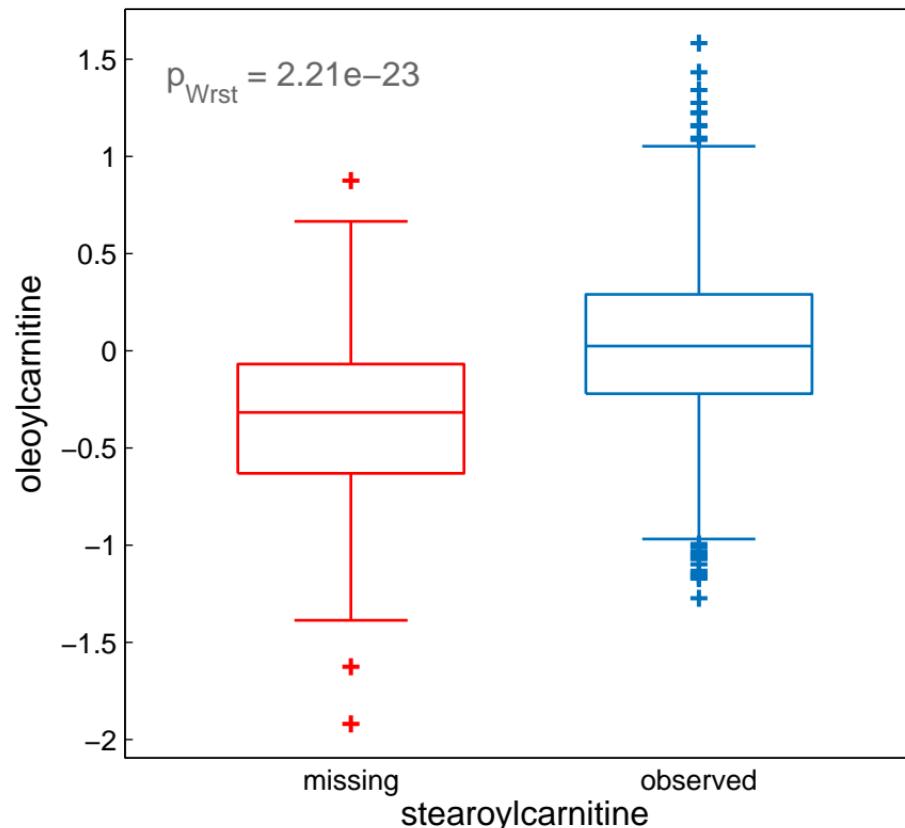


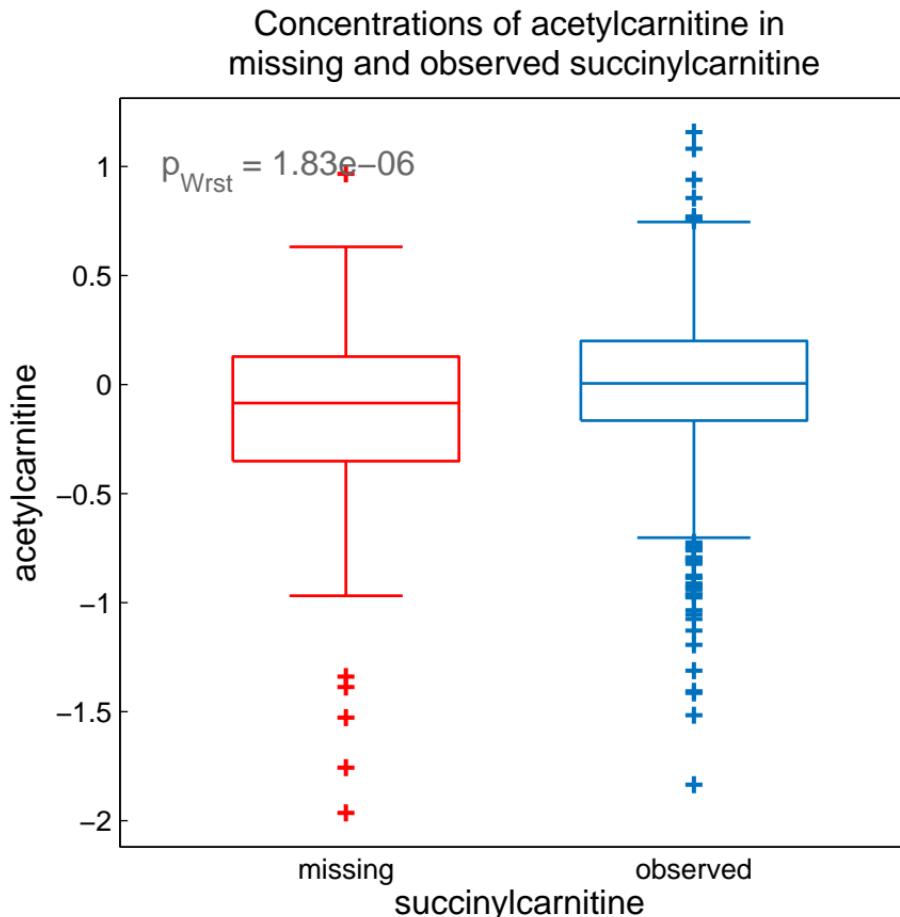
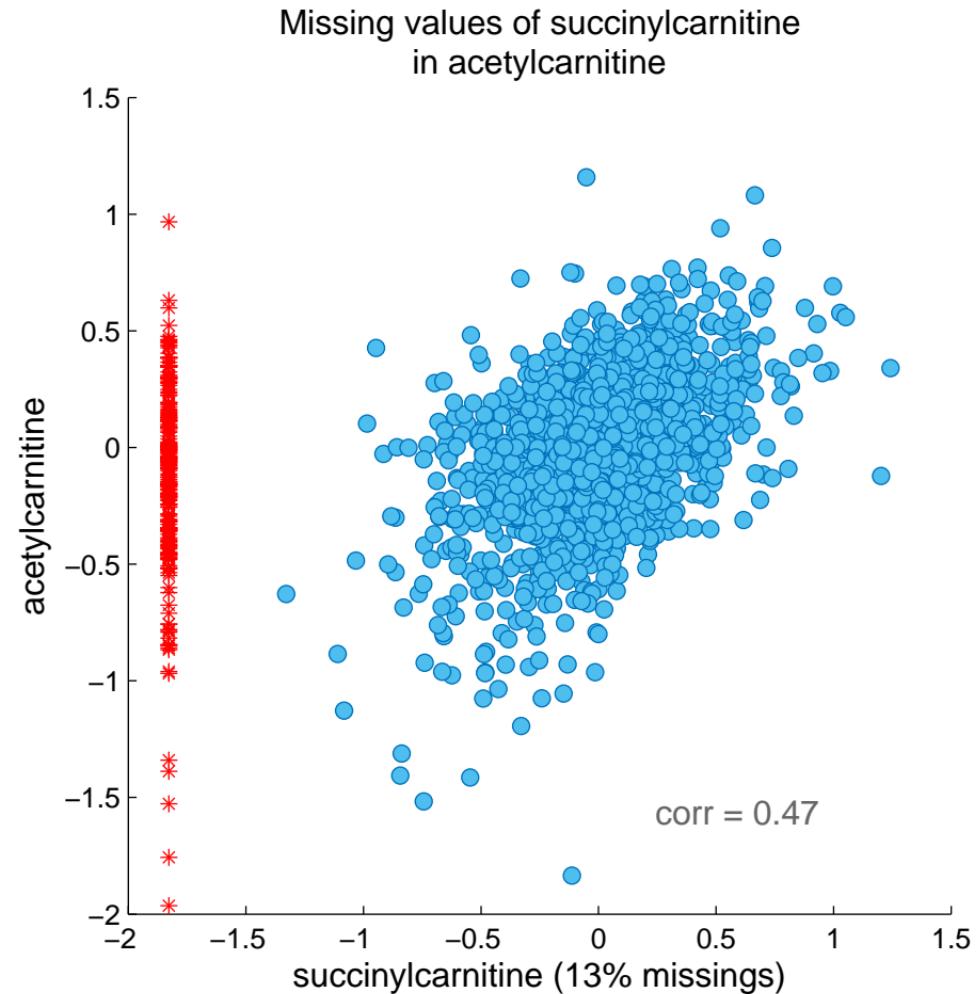


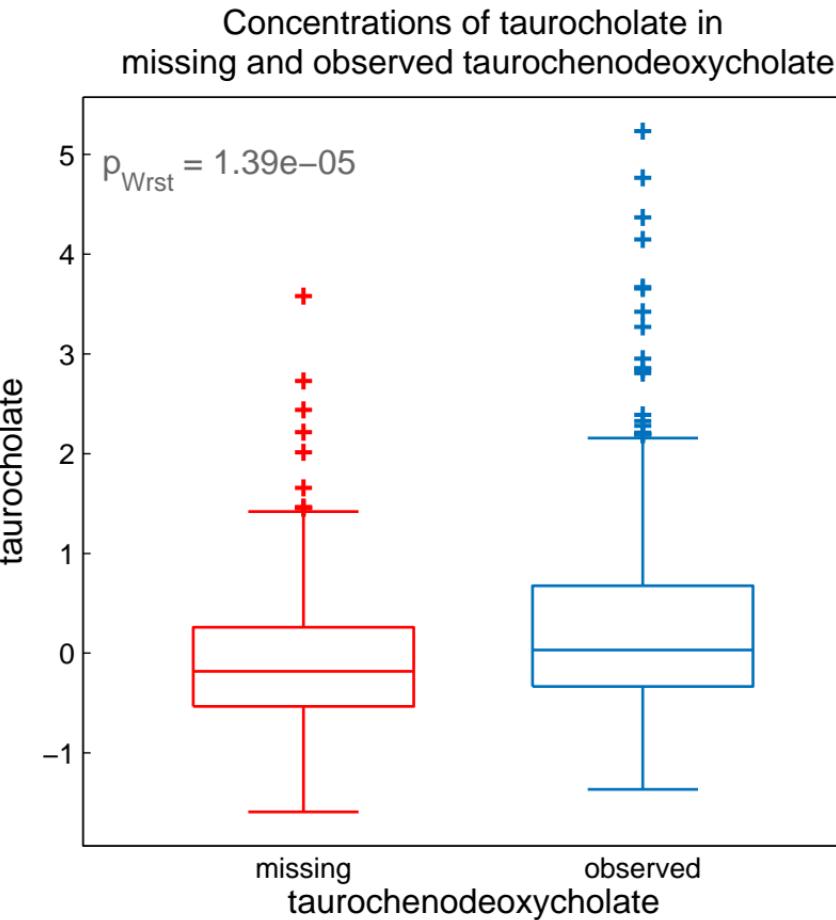
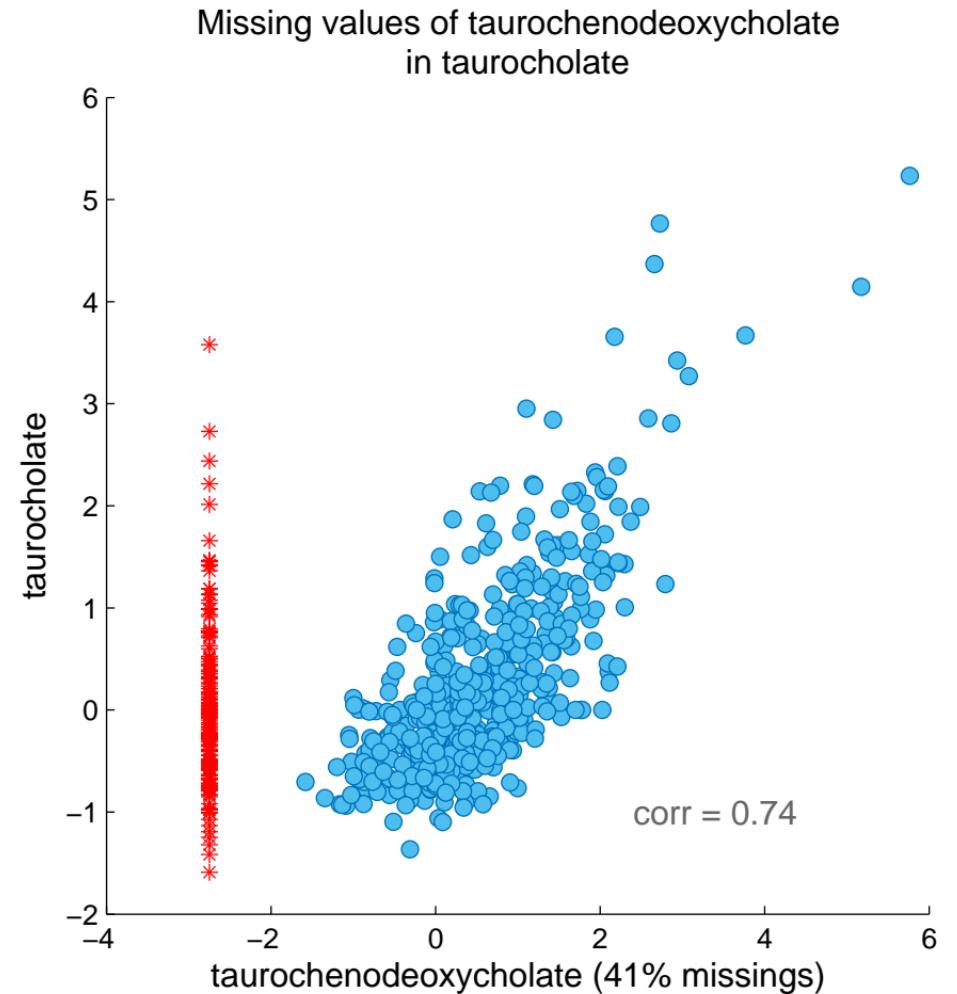
Missing values of stearoylcarnitine
in oleoylcarnitine

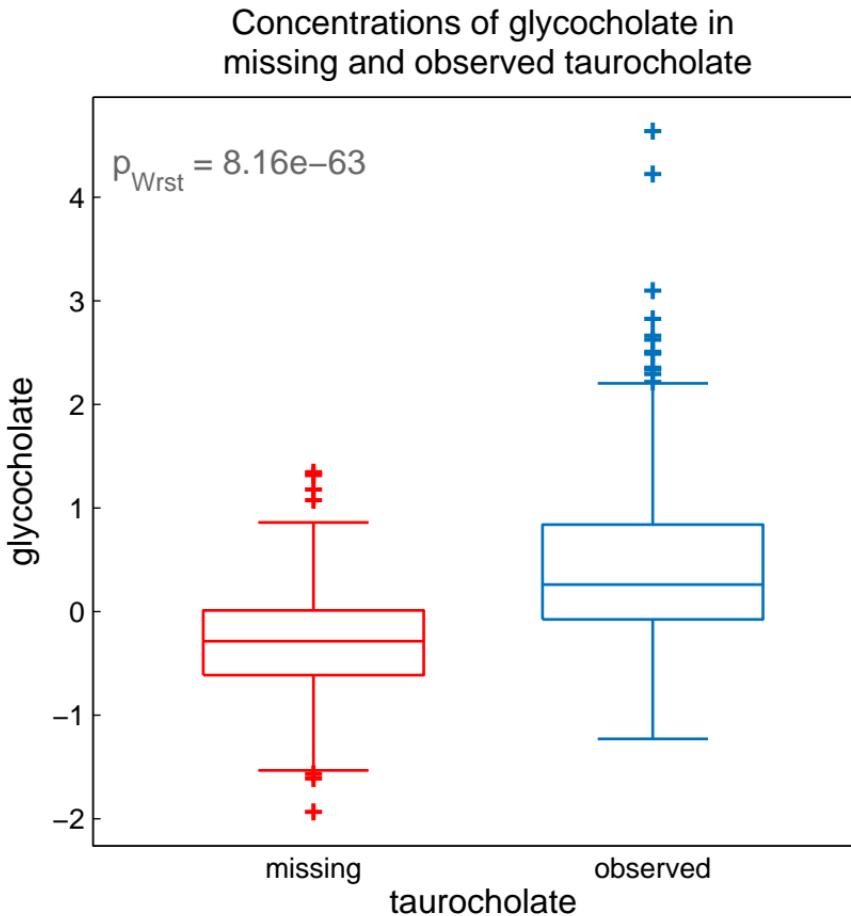
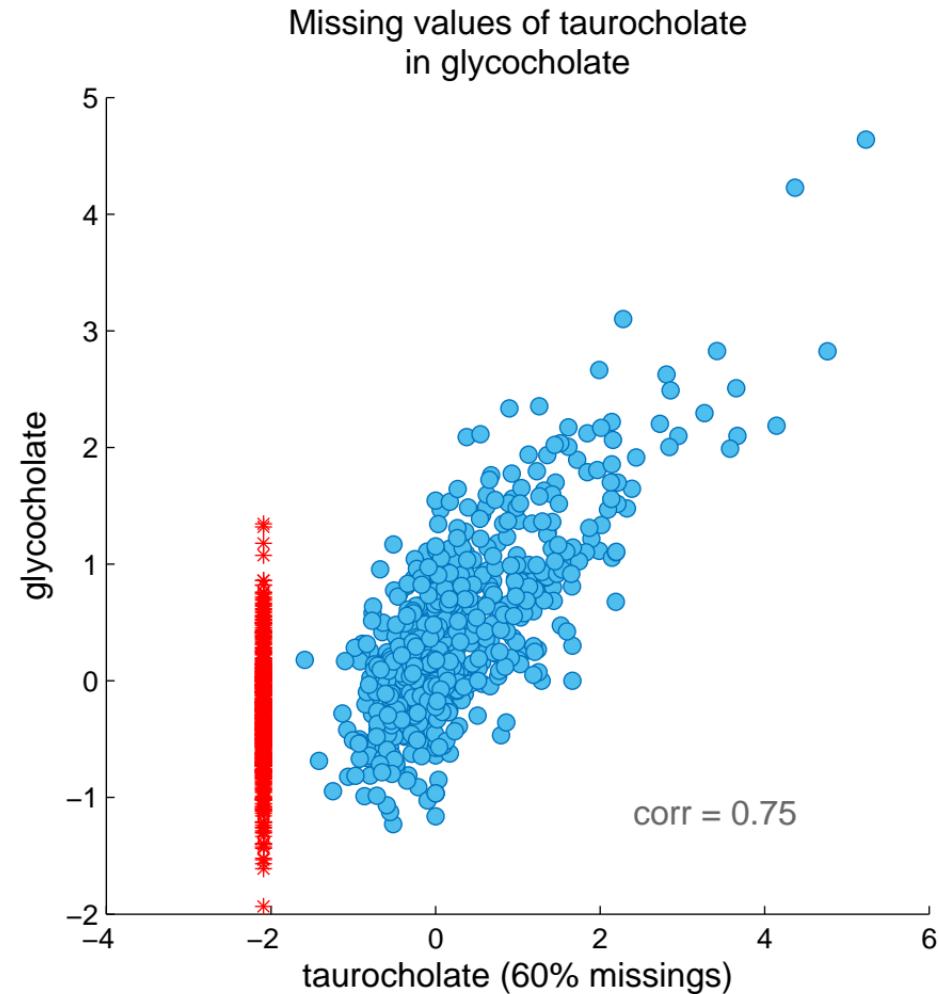


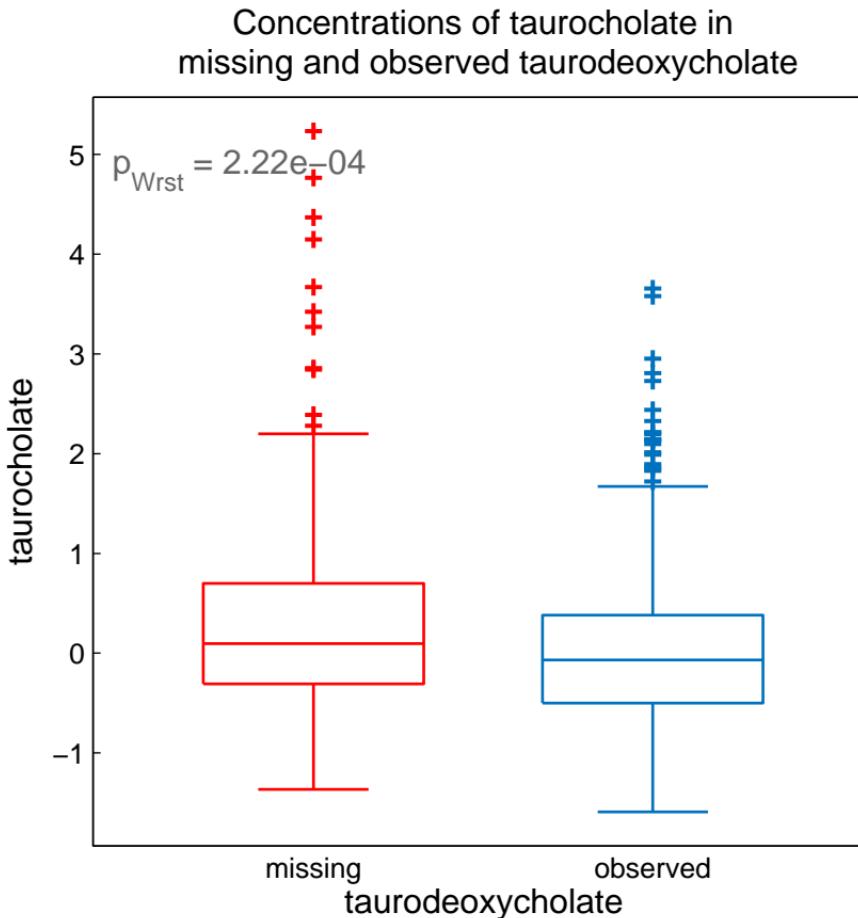
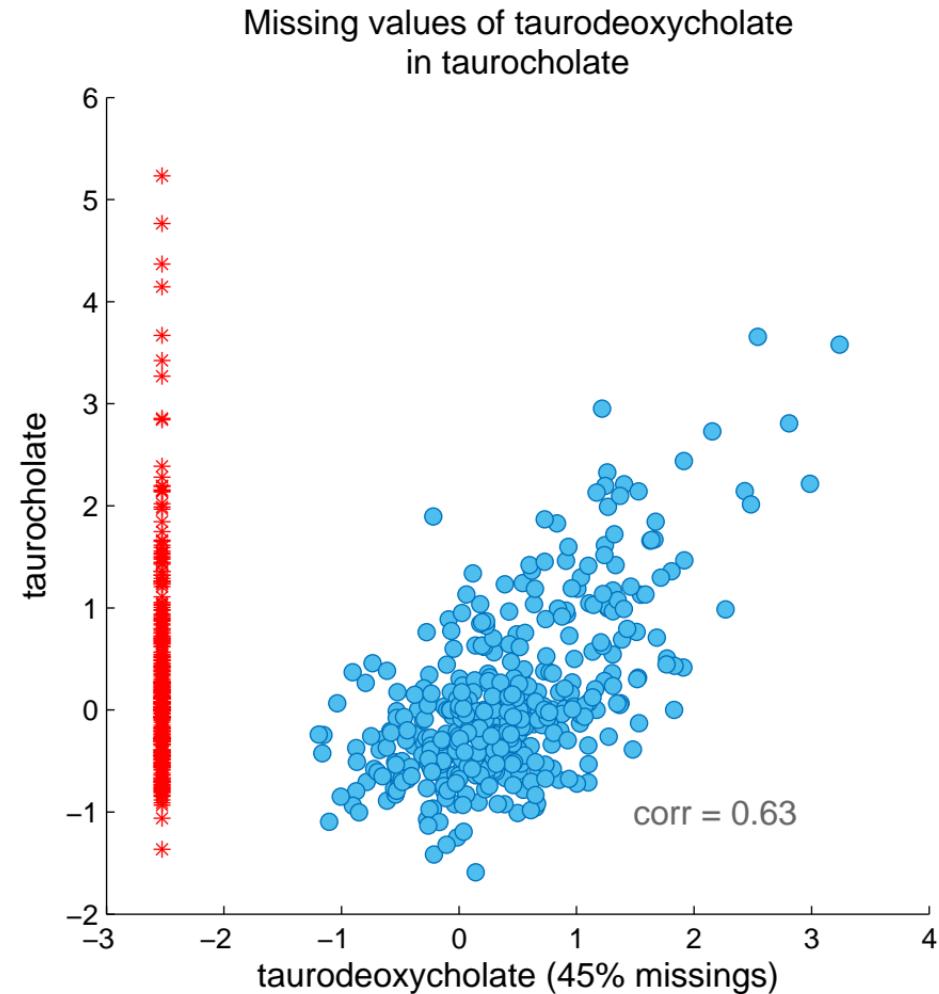
Concentrations of oleoylcarnitine in
missing and observed stearoylcarnitine

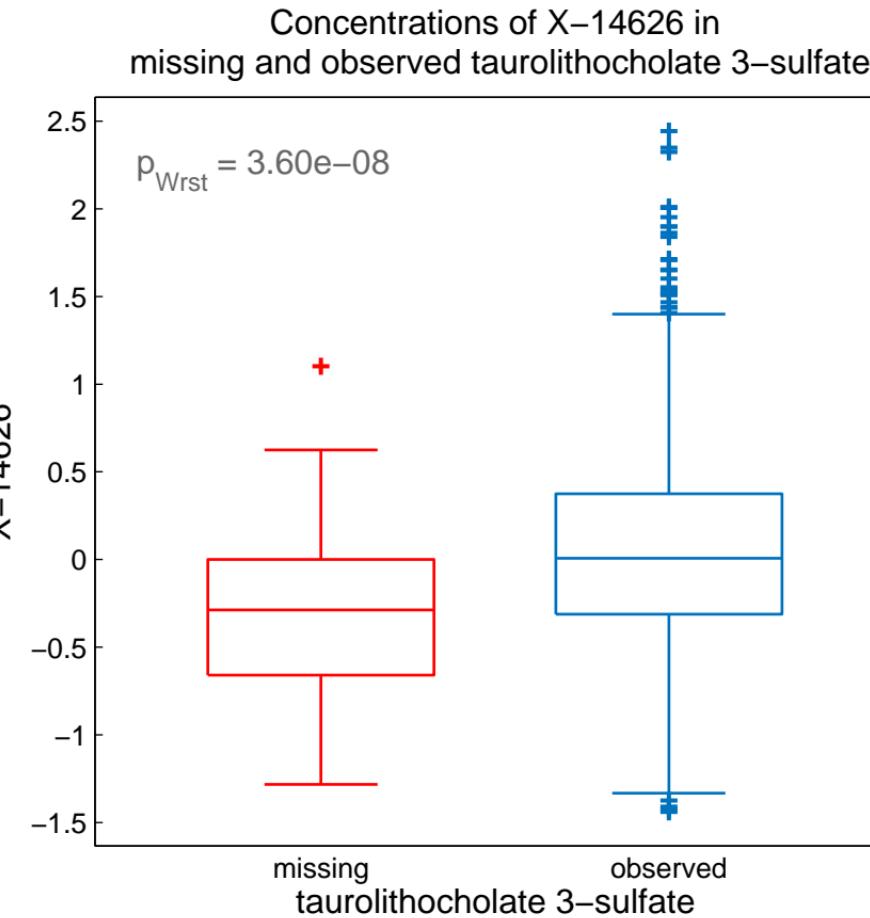
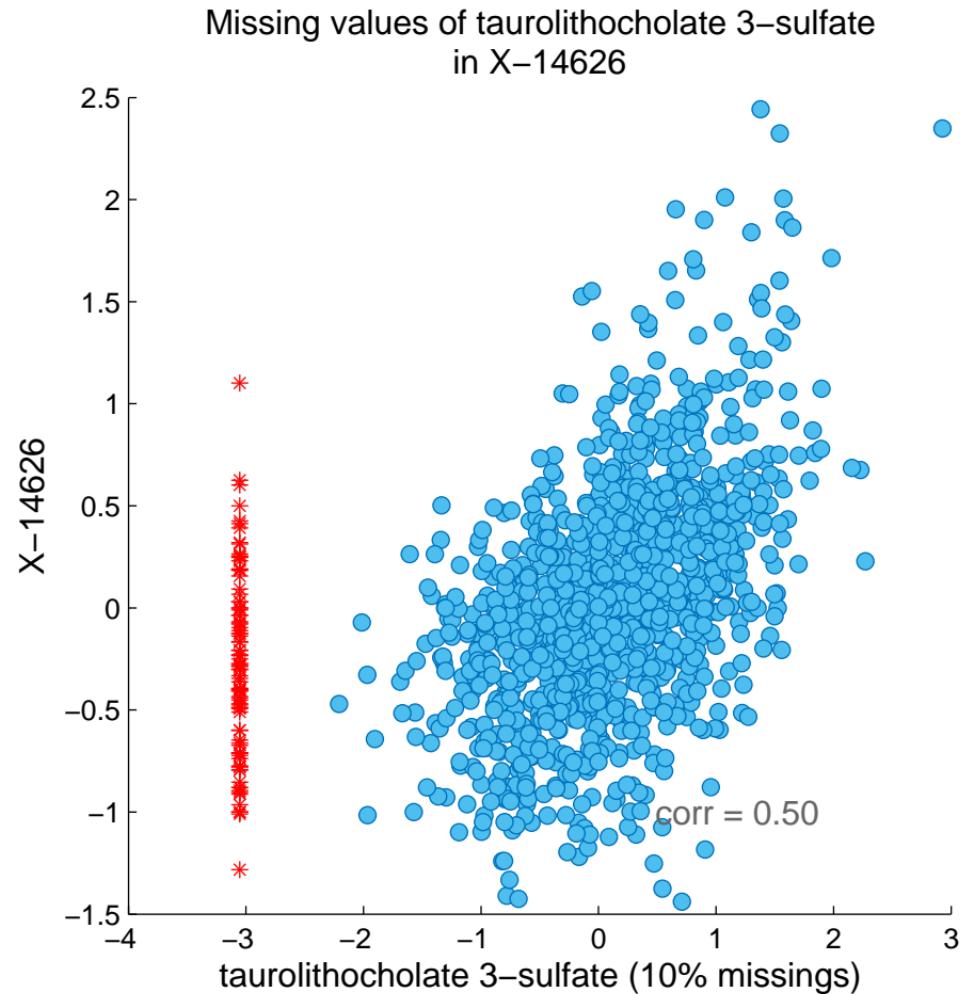


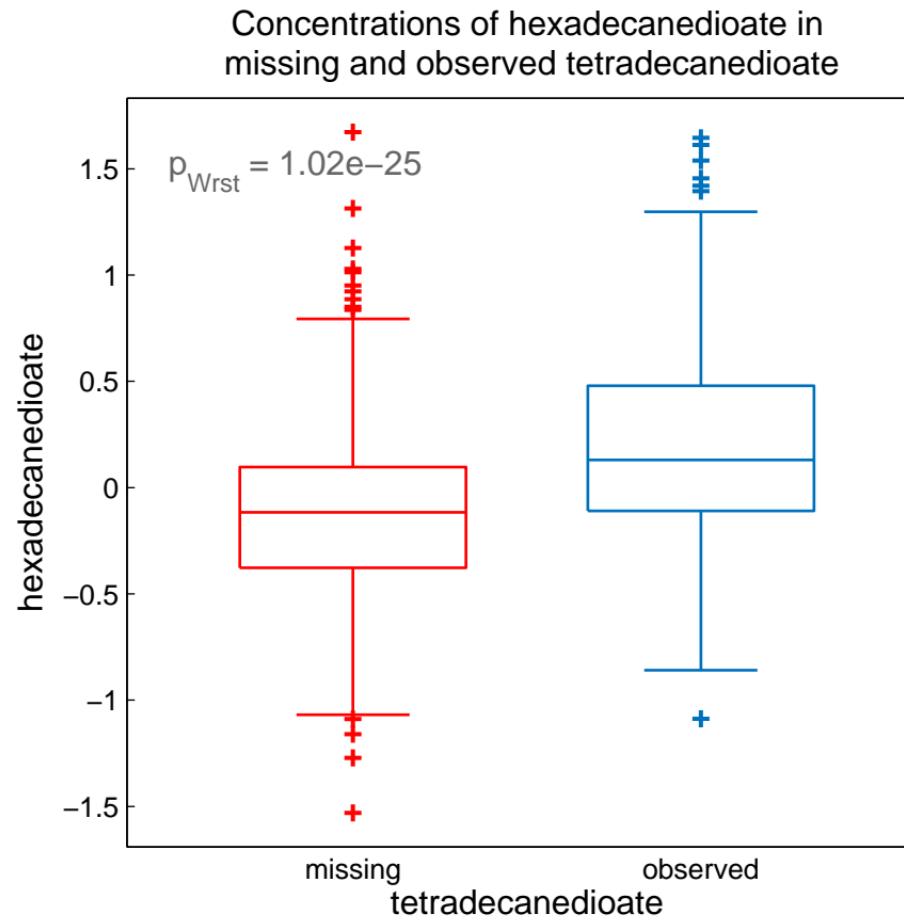
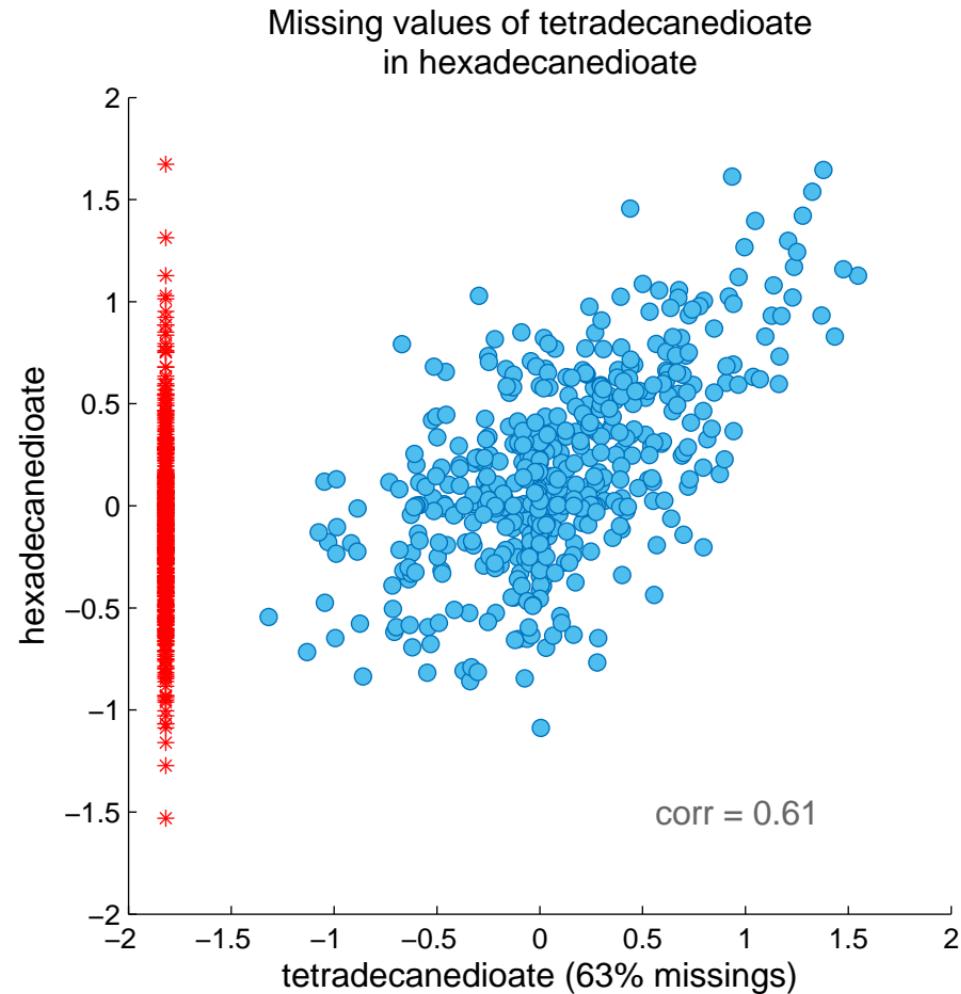


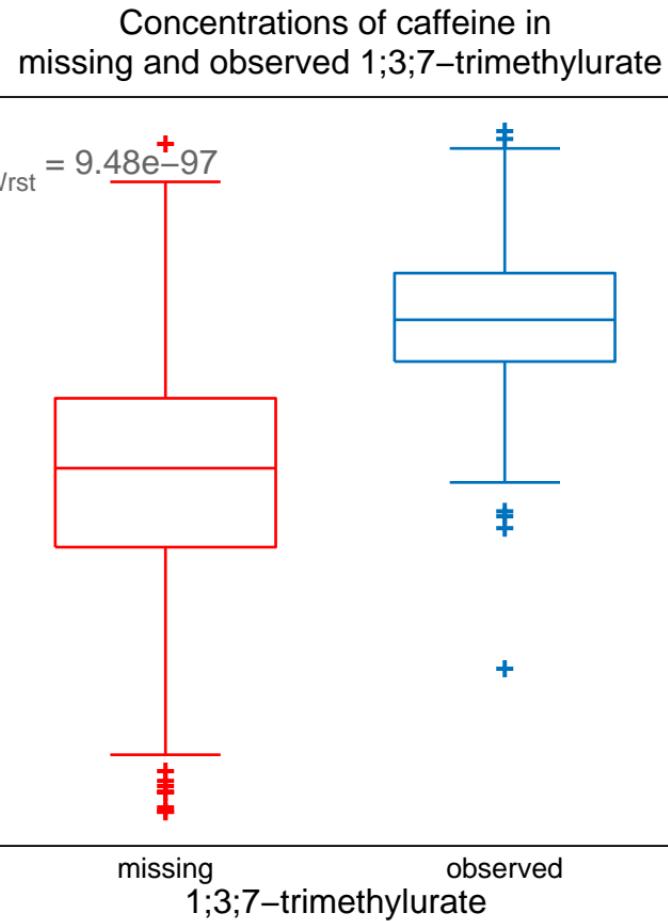
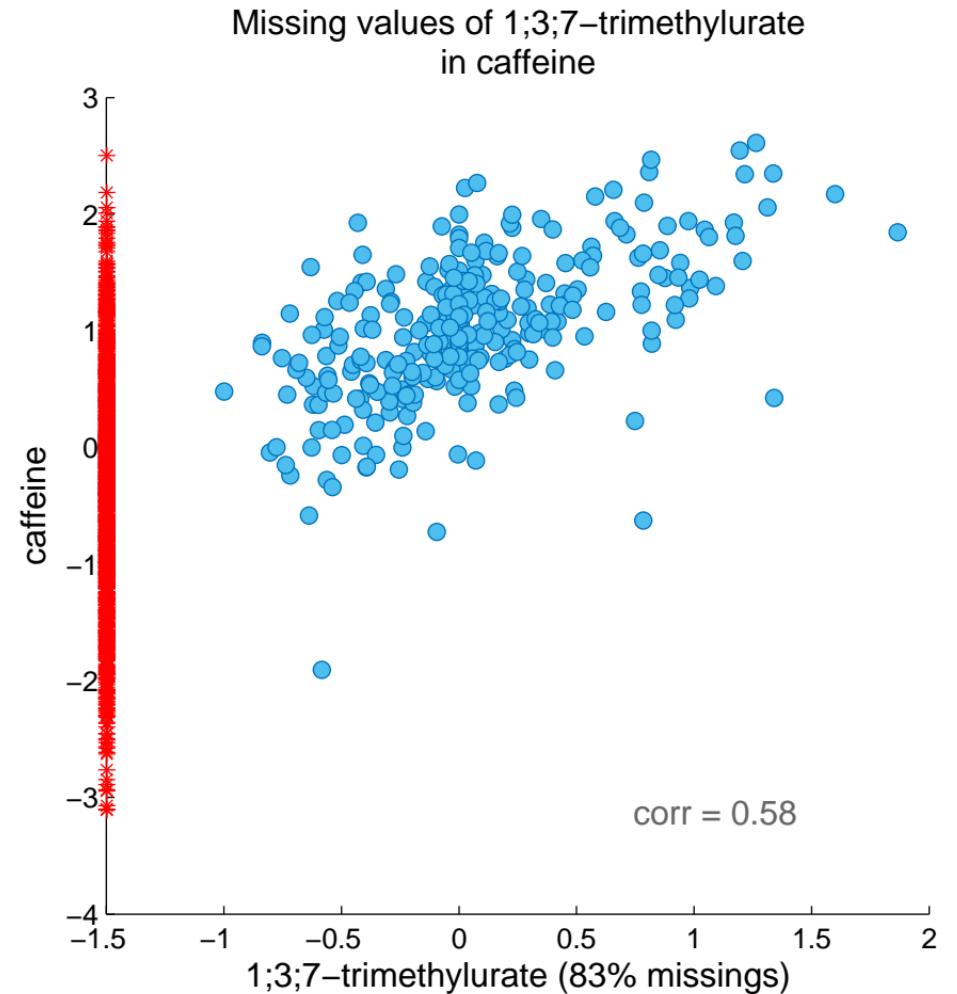


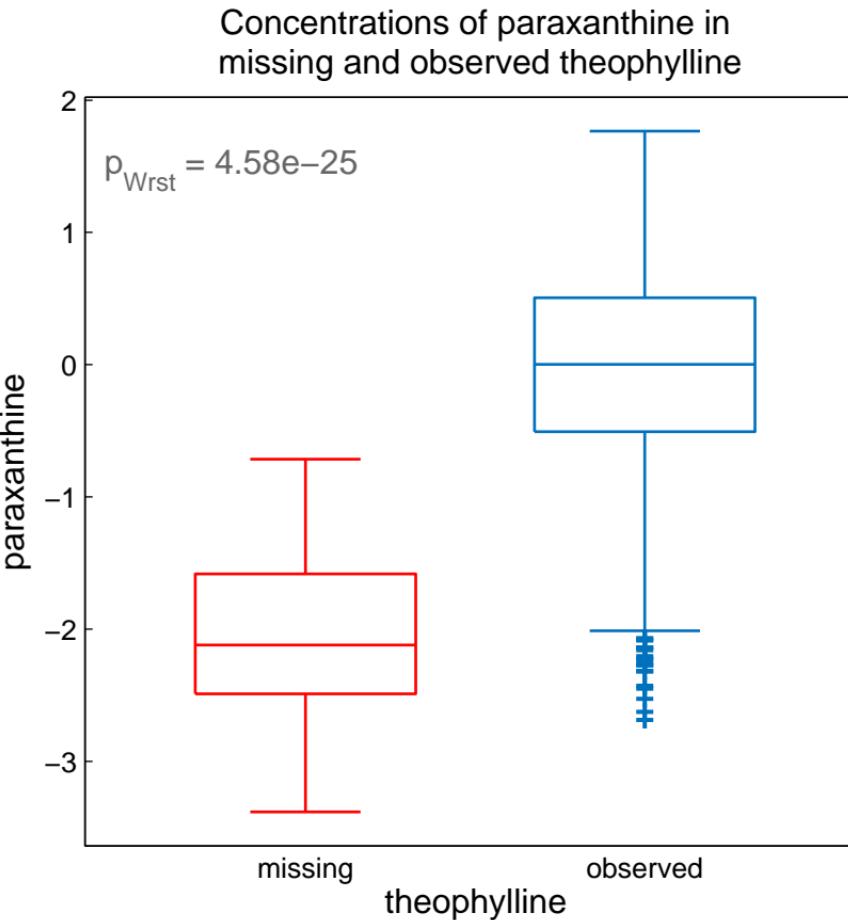
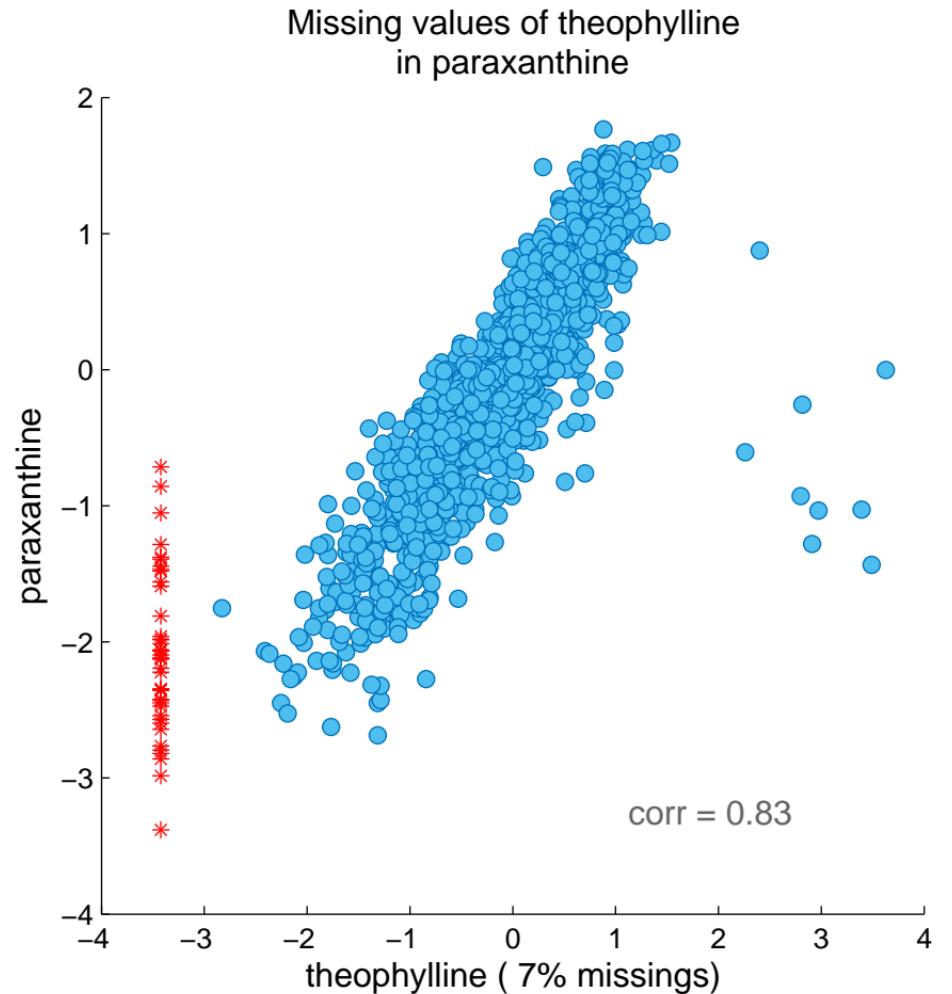


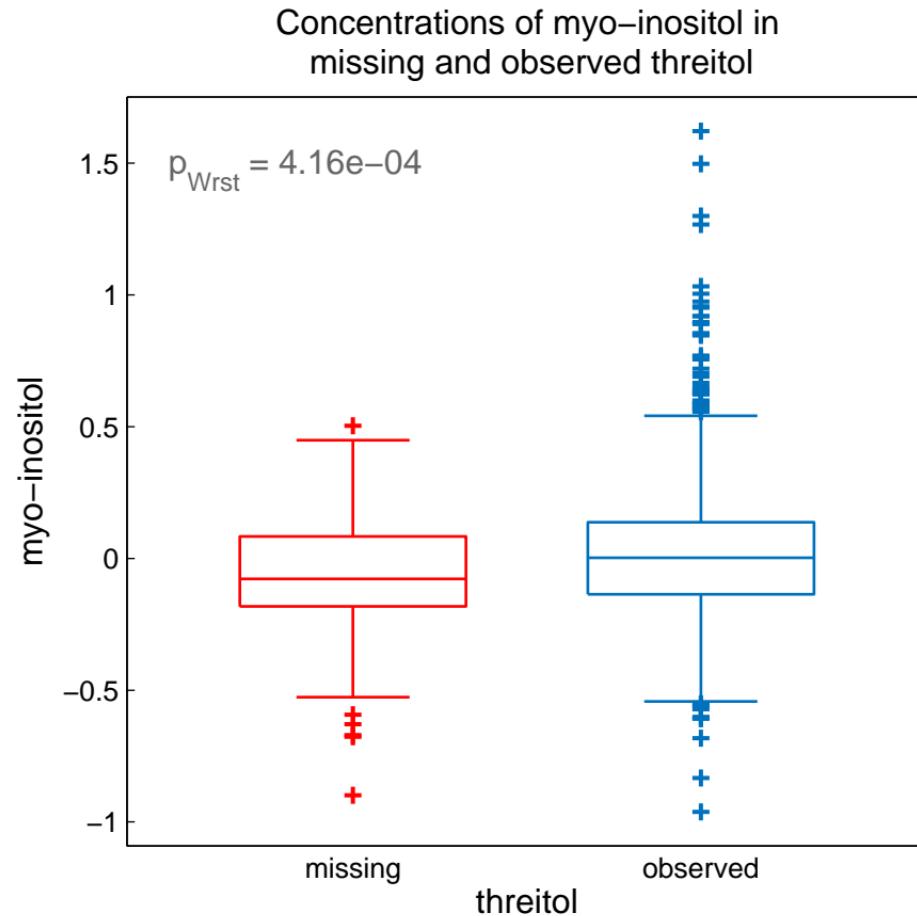
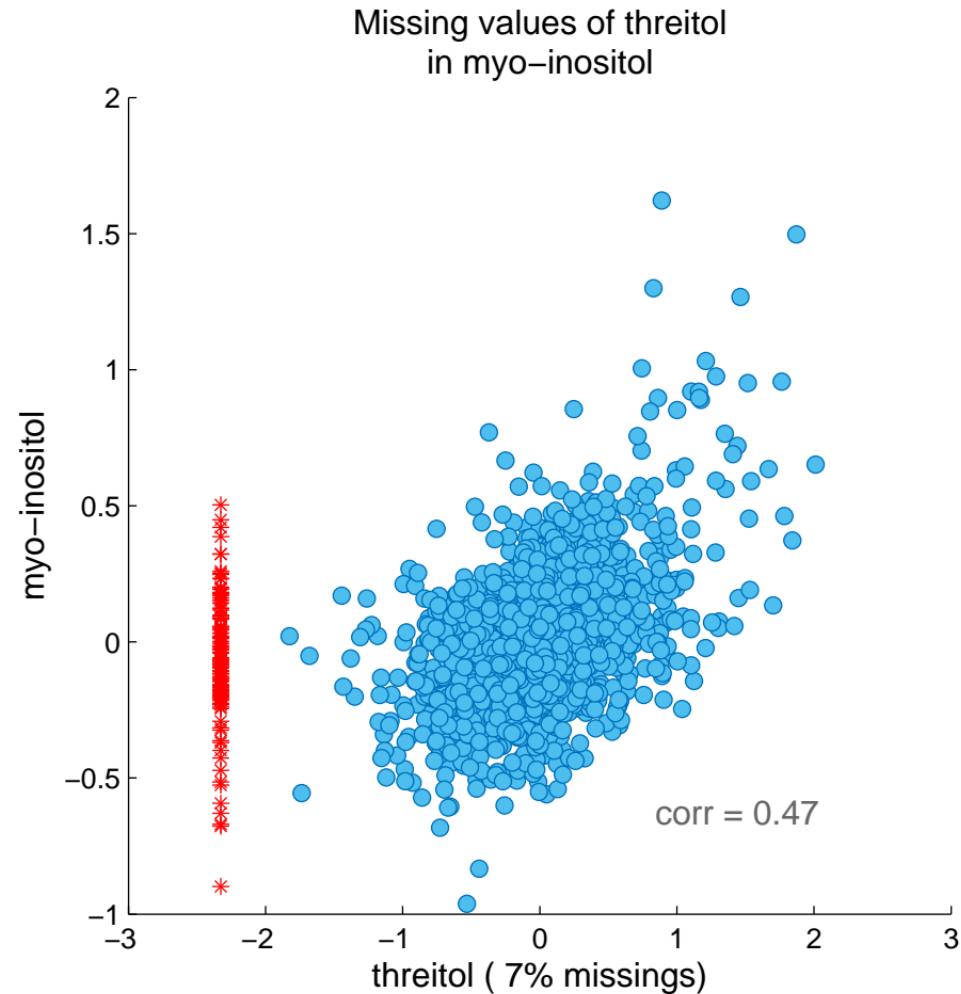


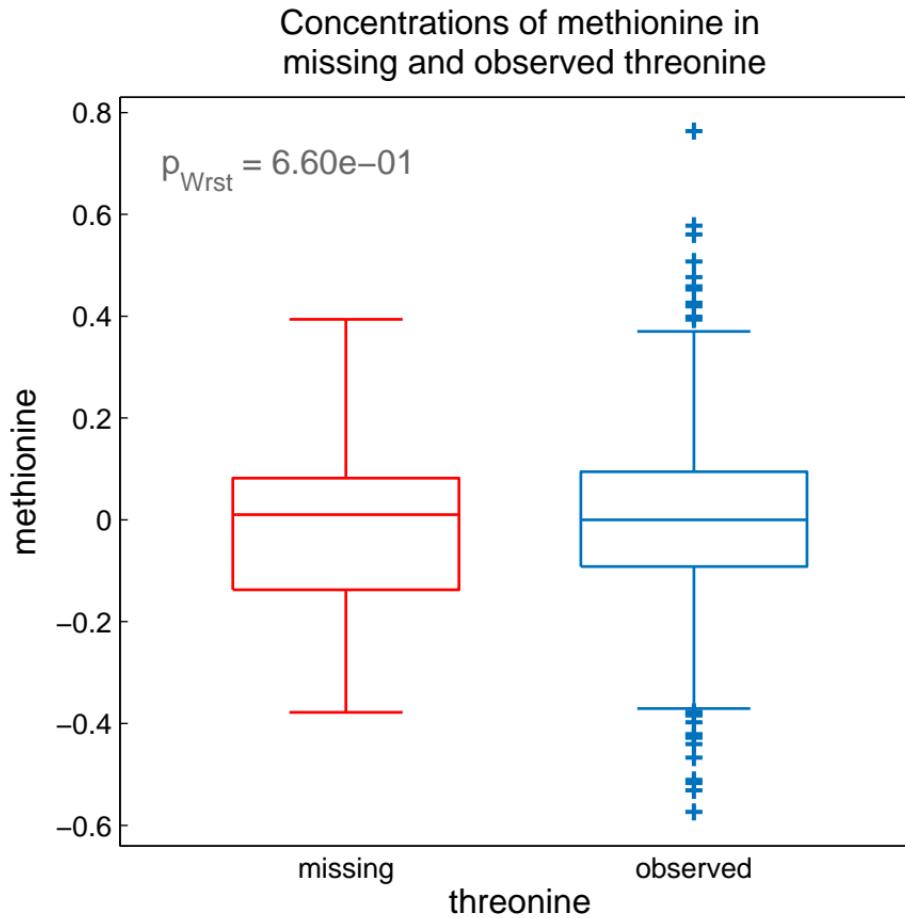
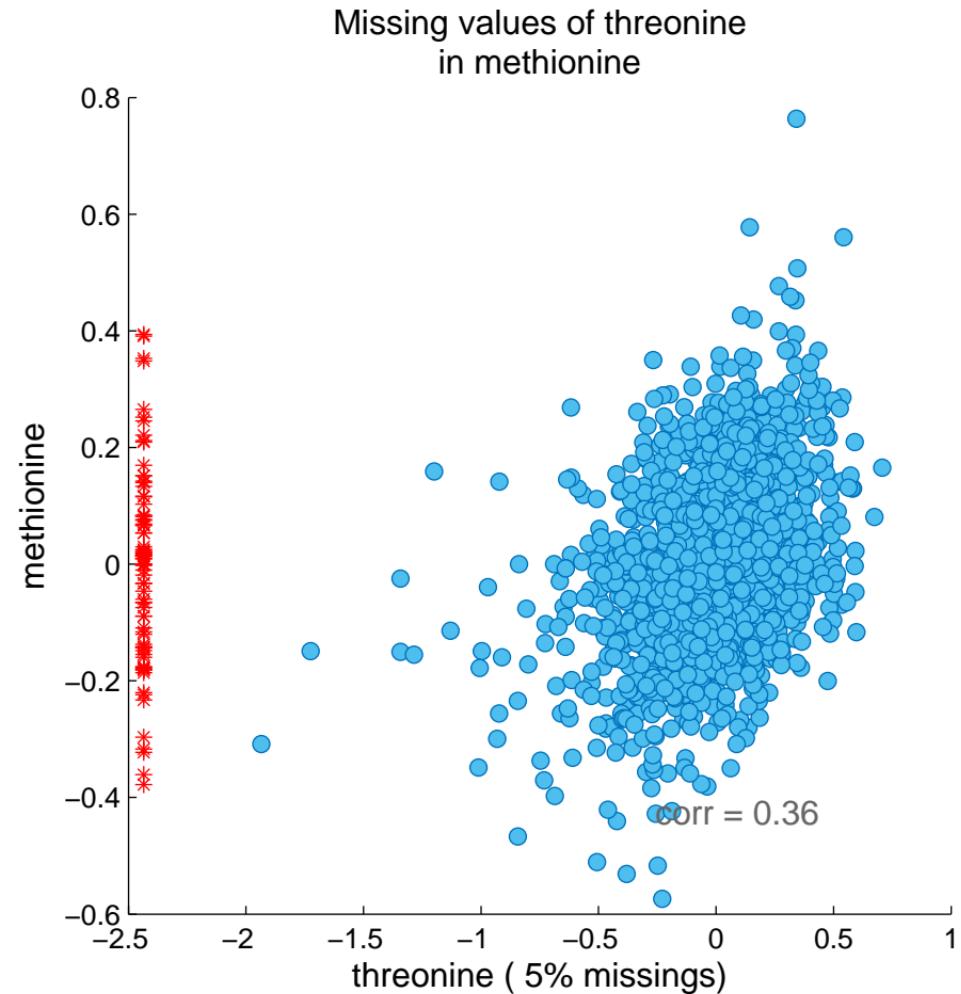


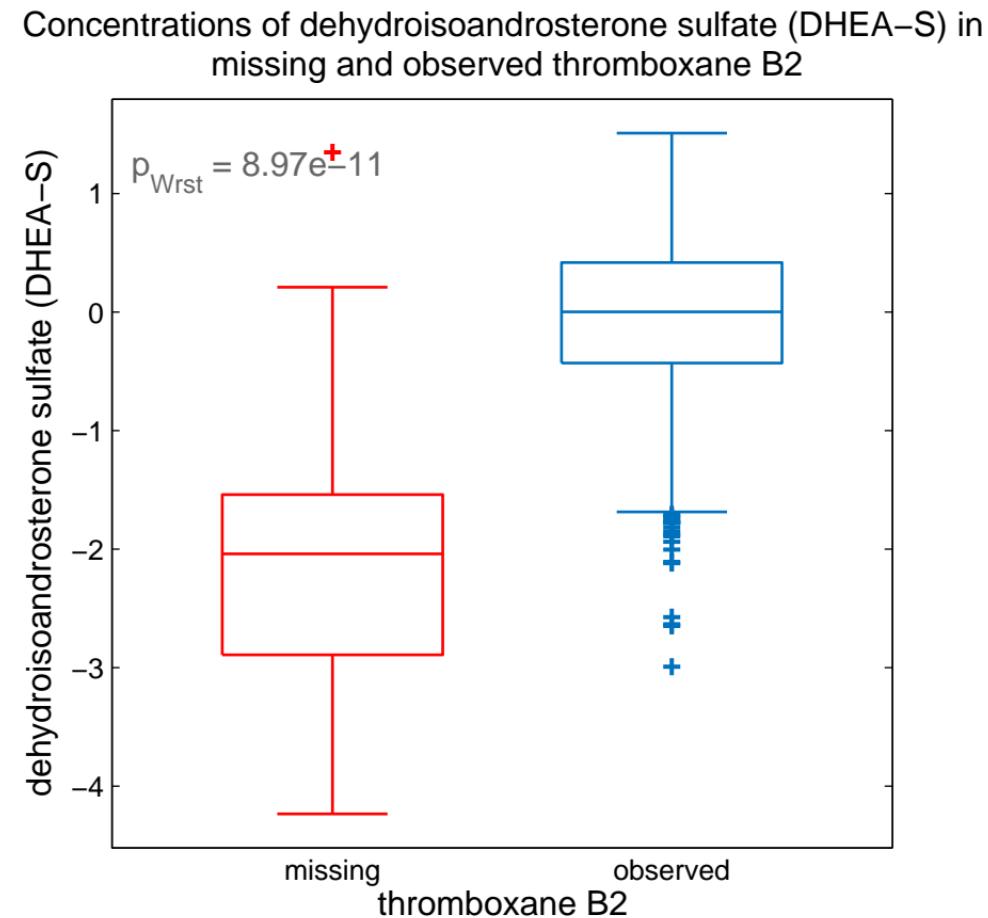
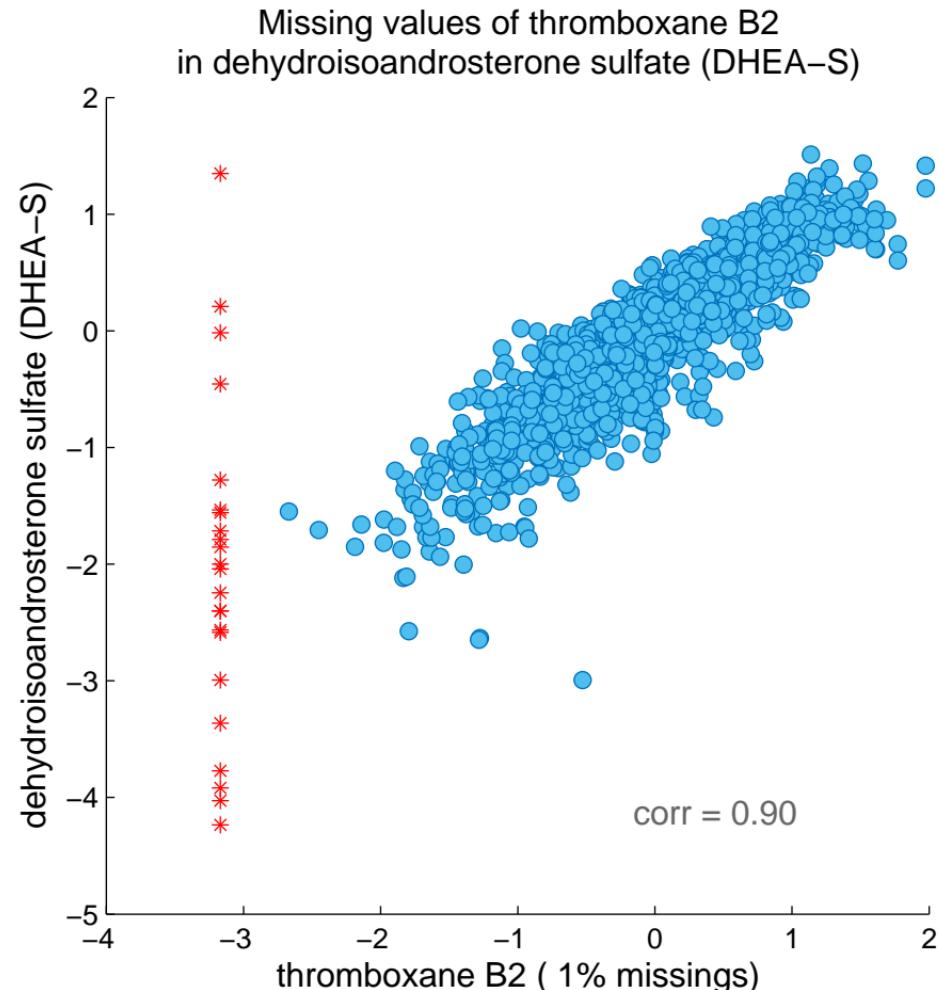


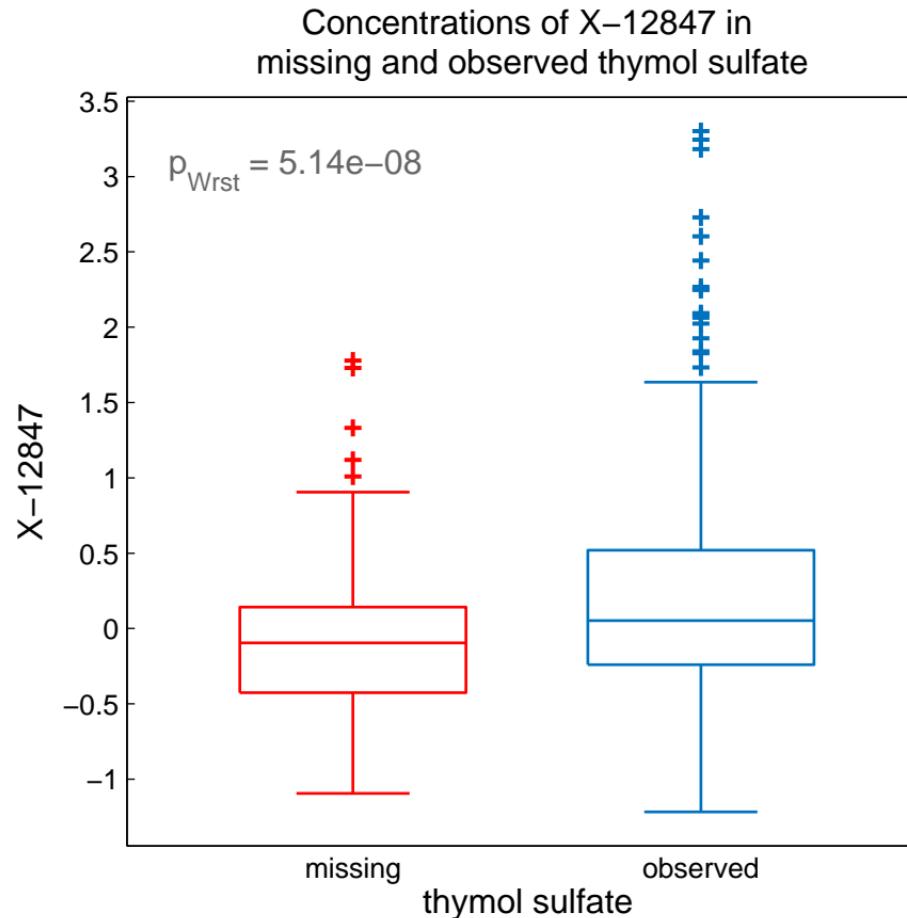
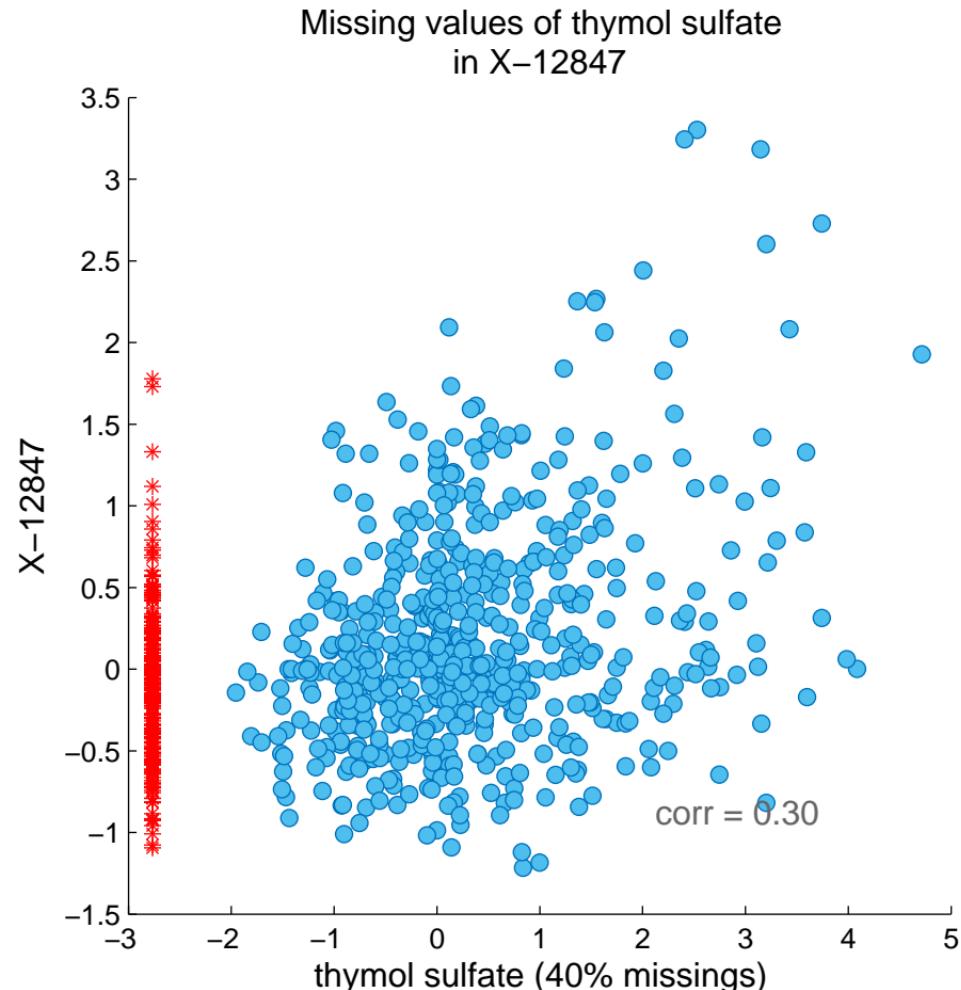


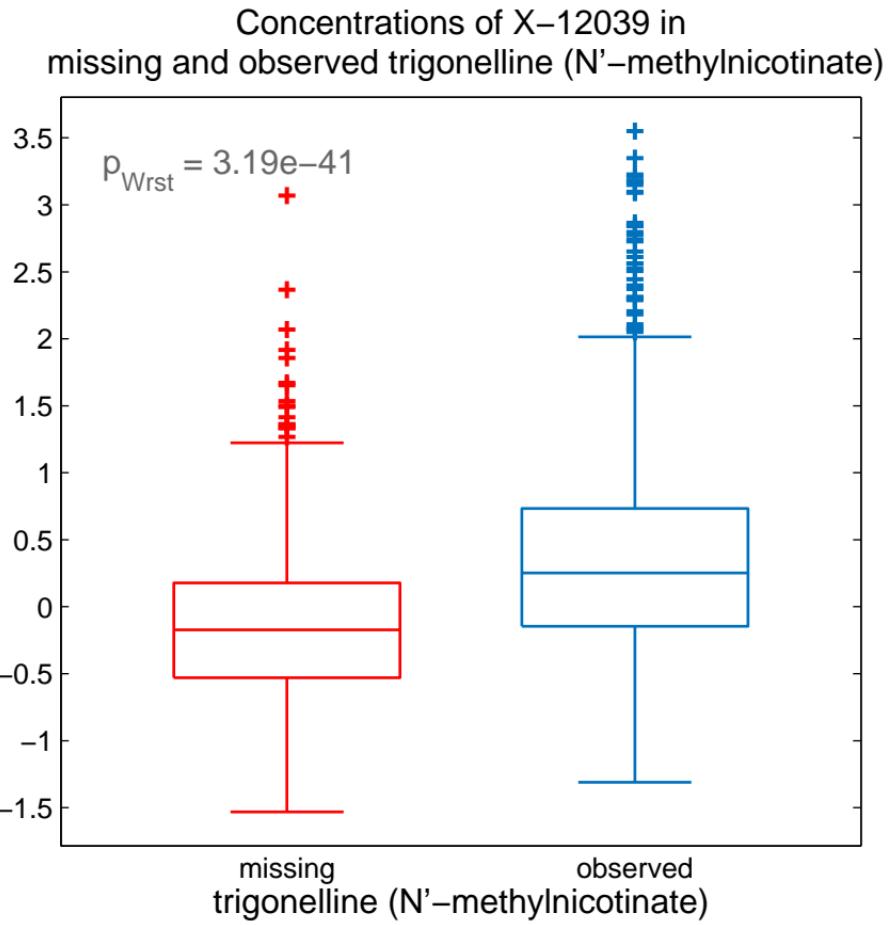
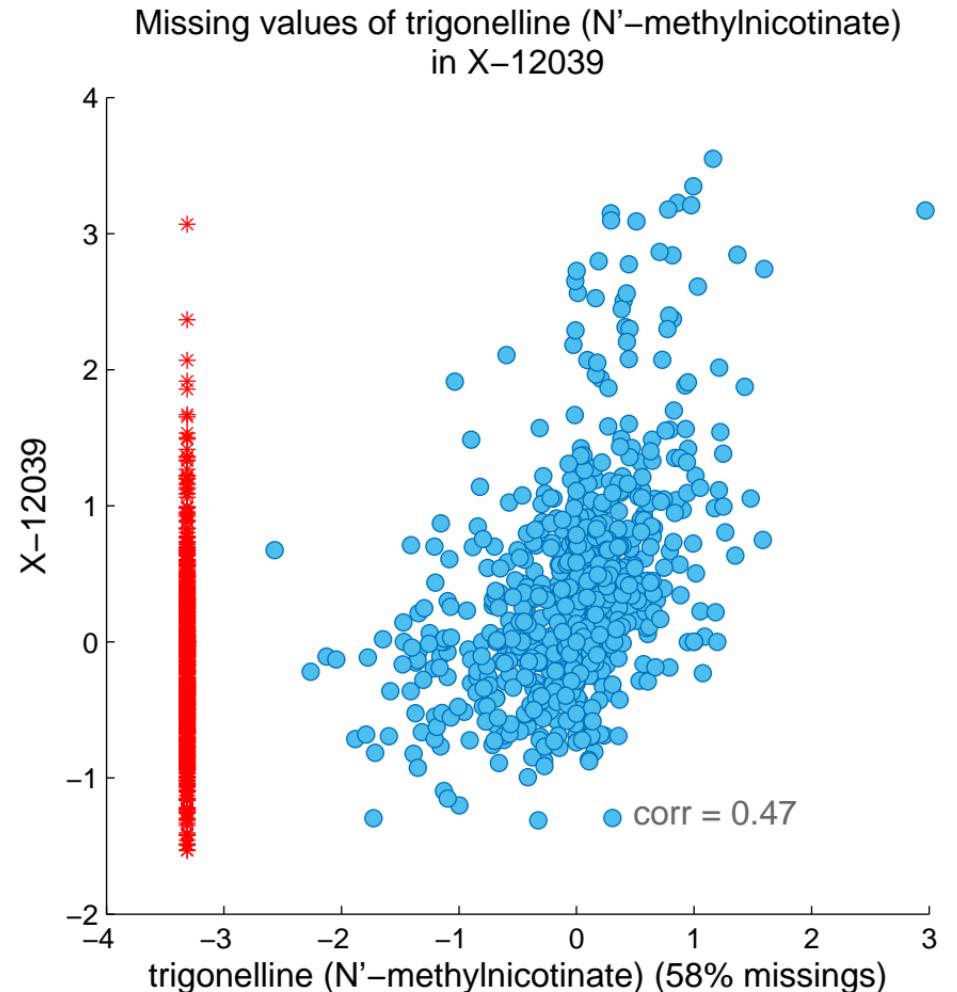


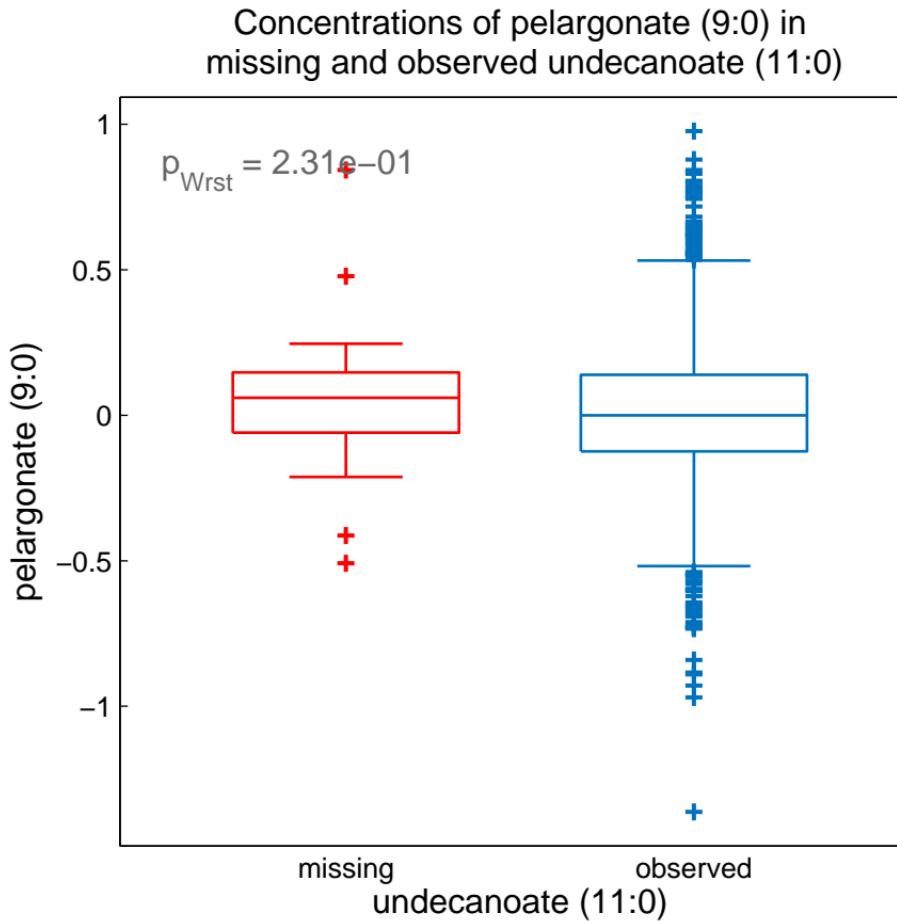
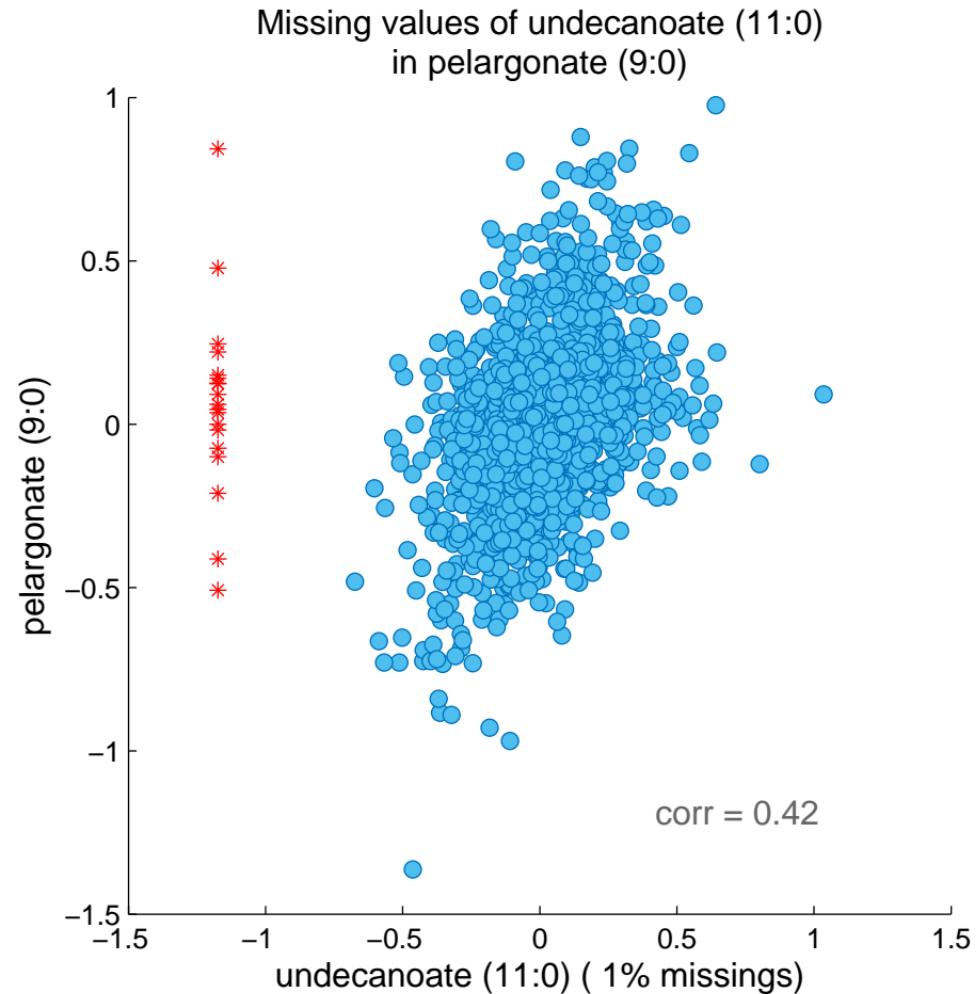


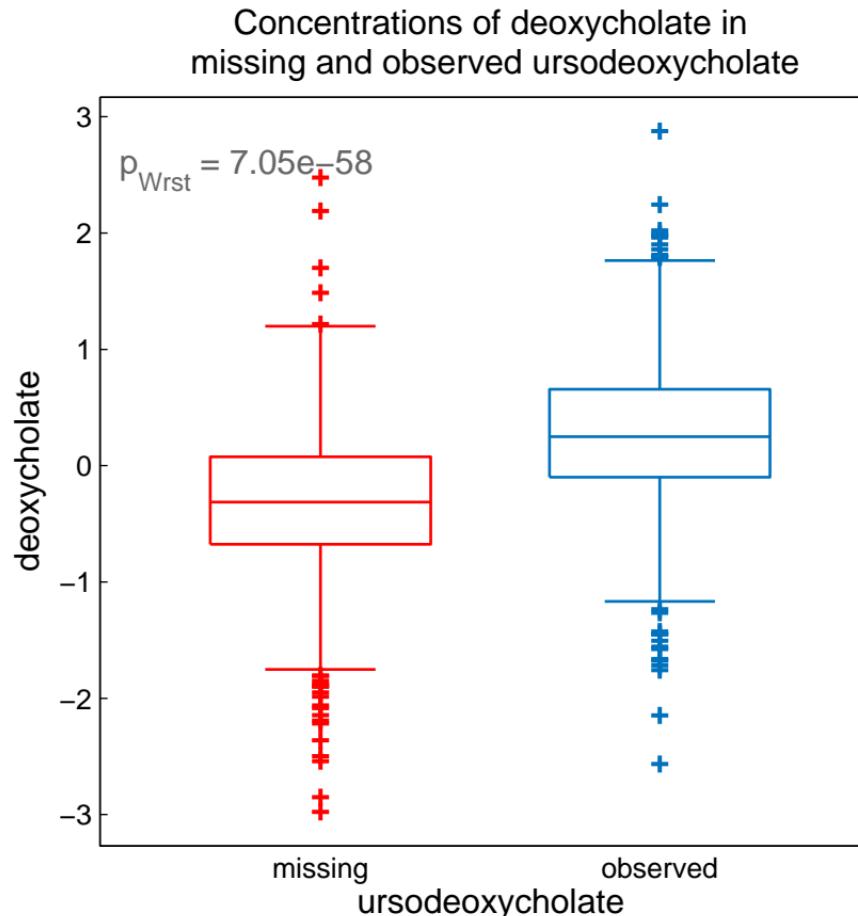
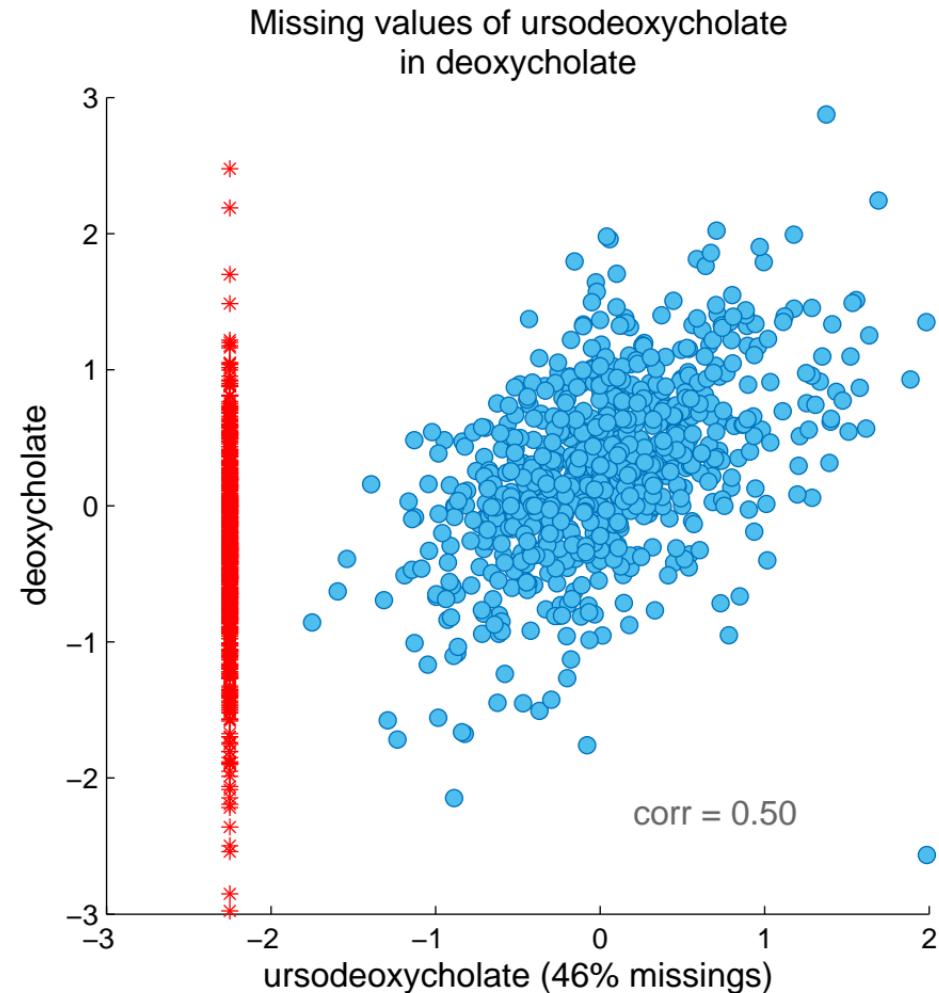


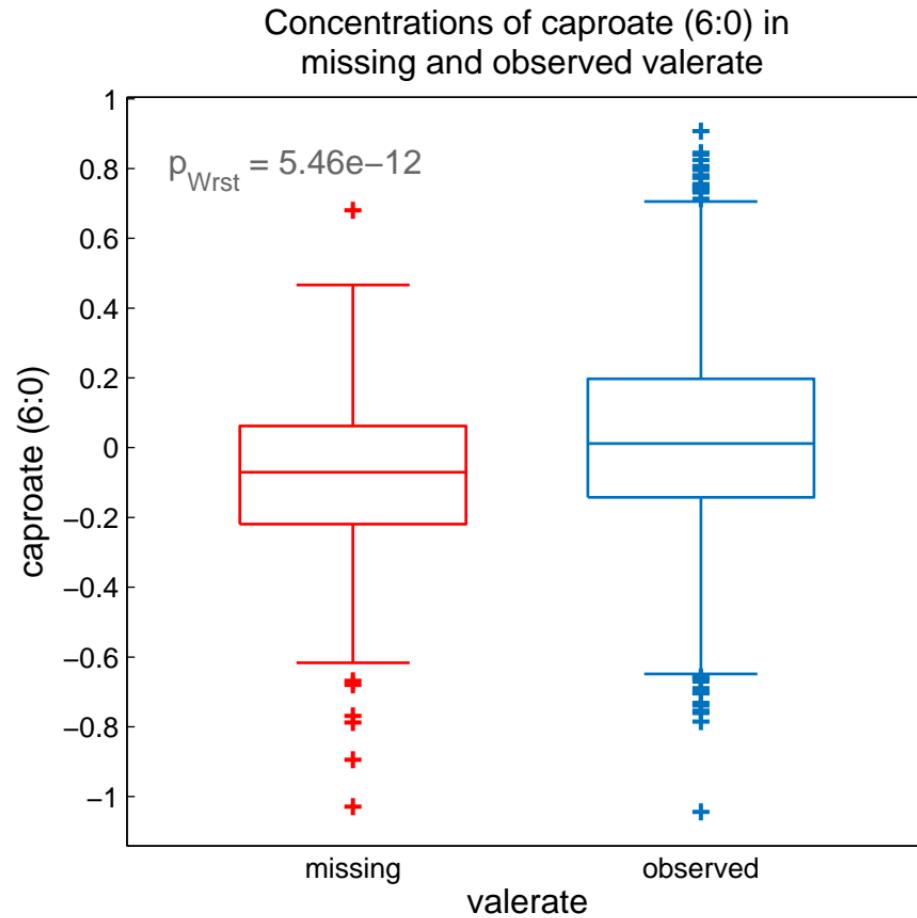
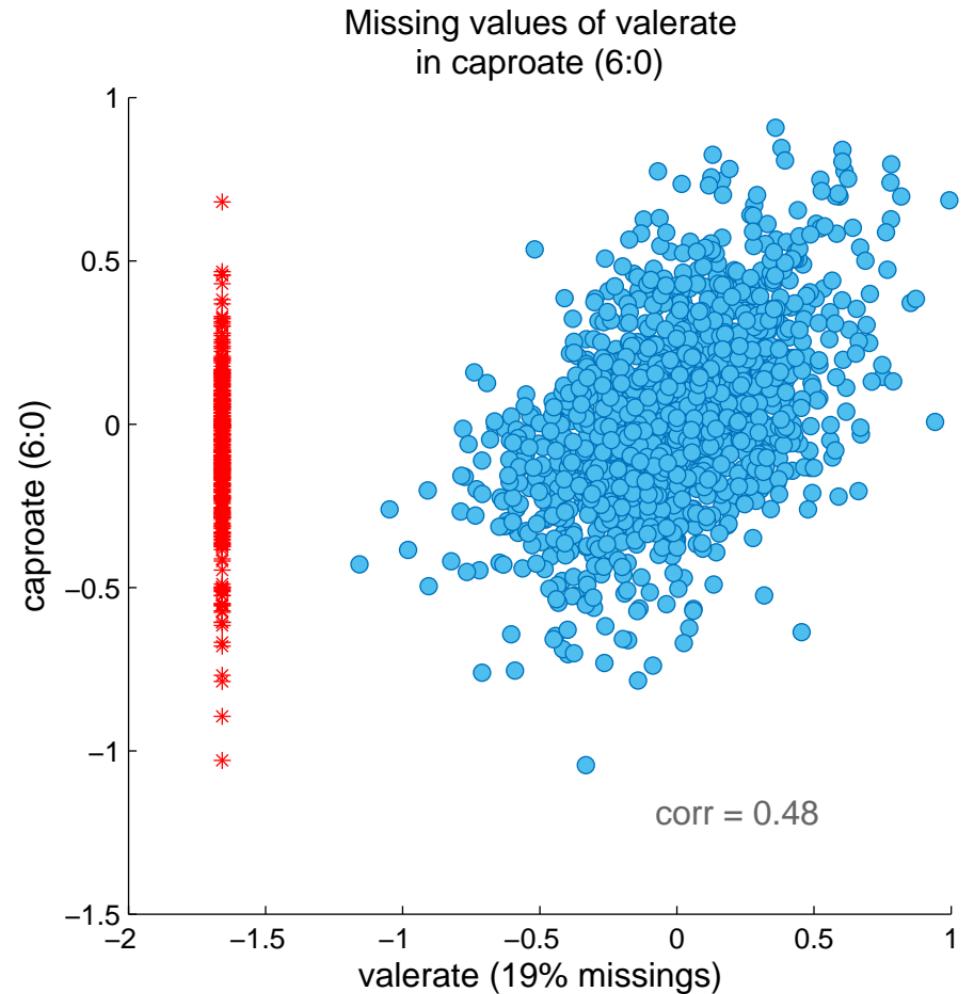


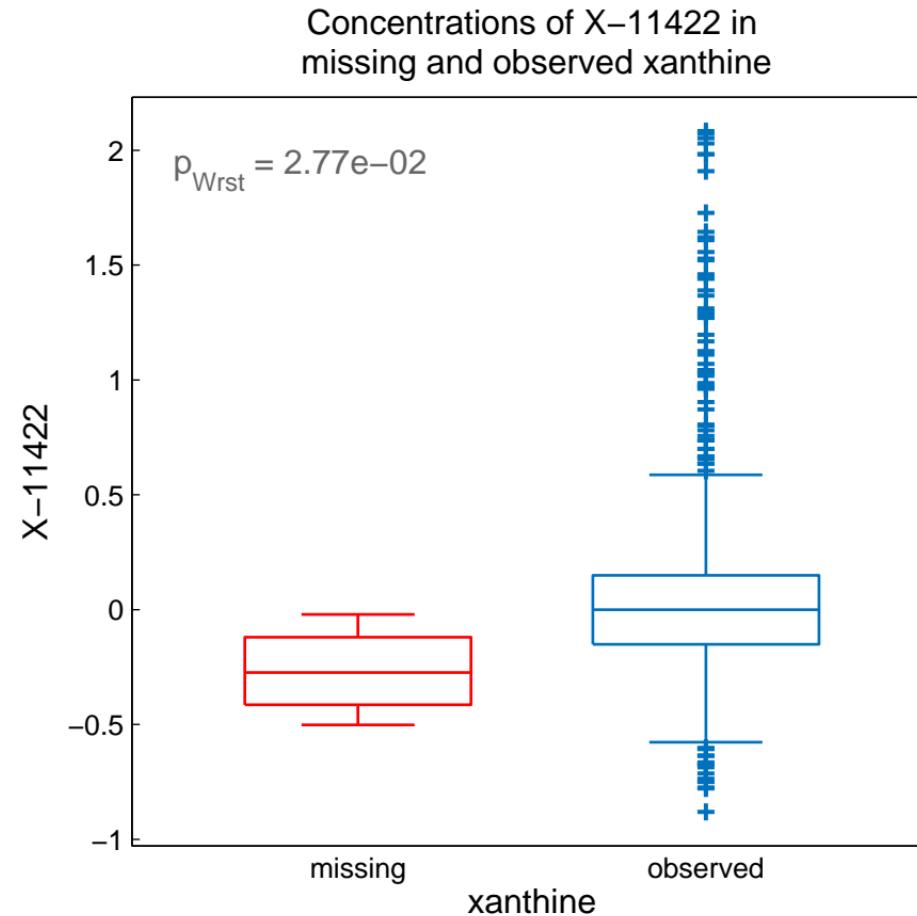
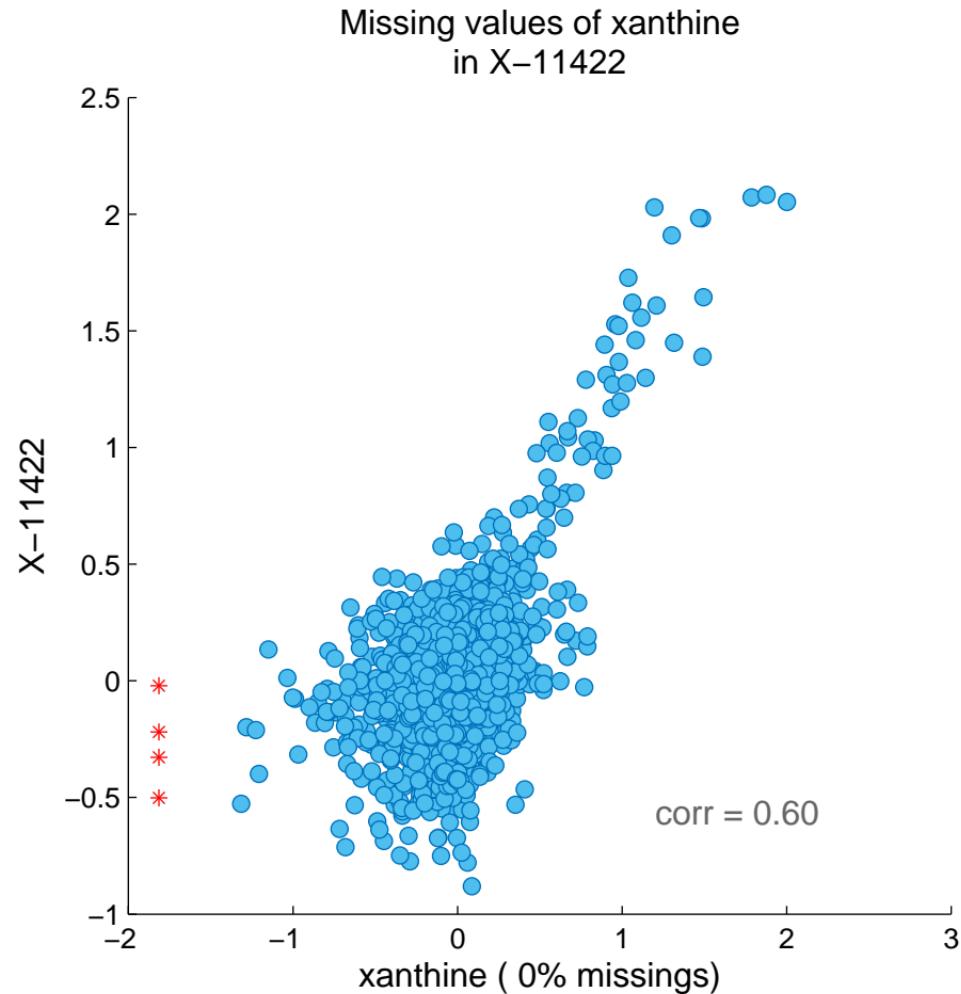


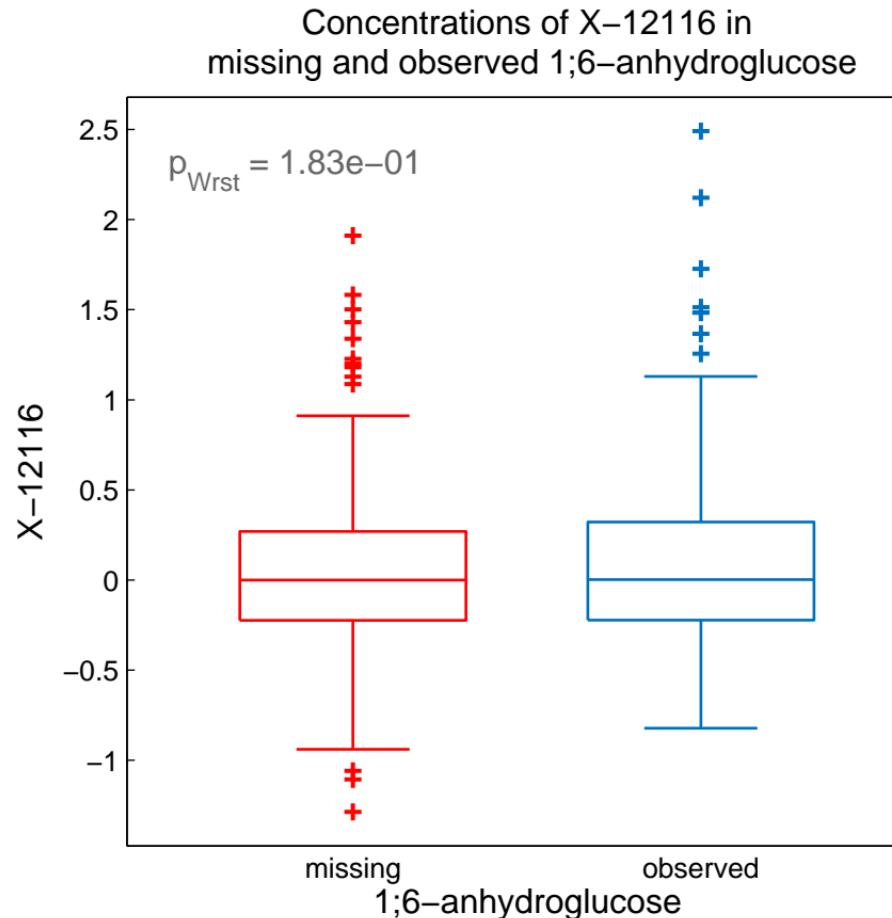
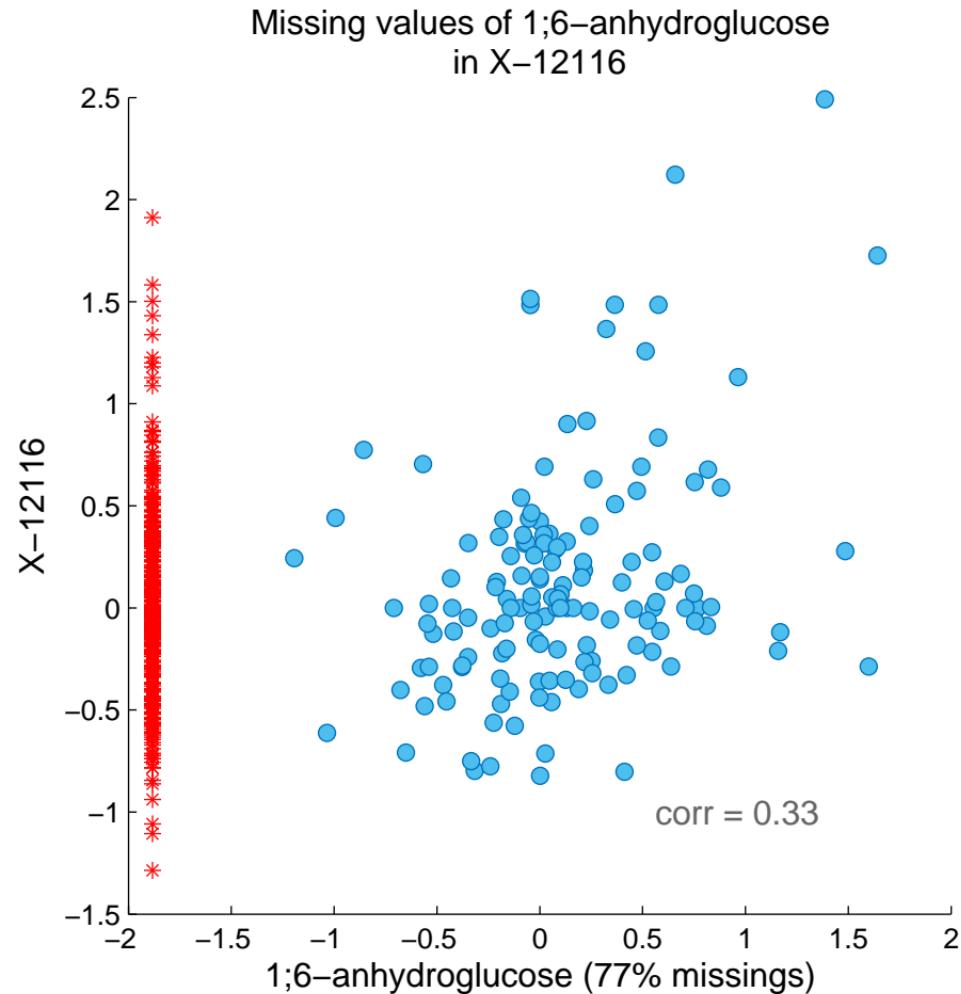


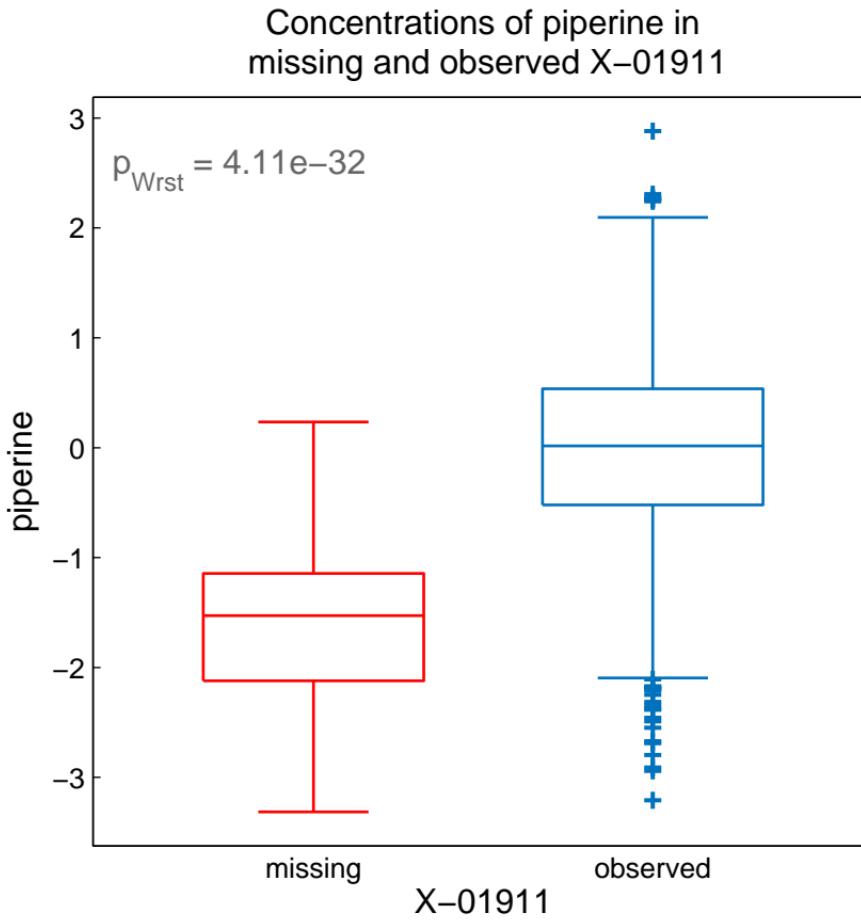
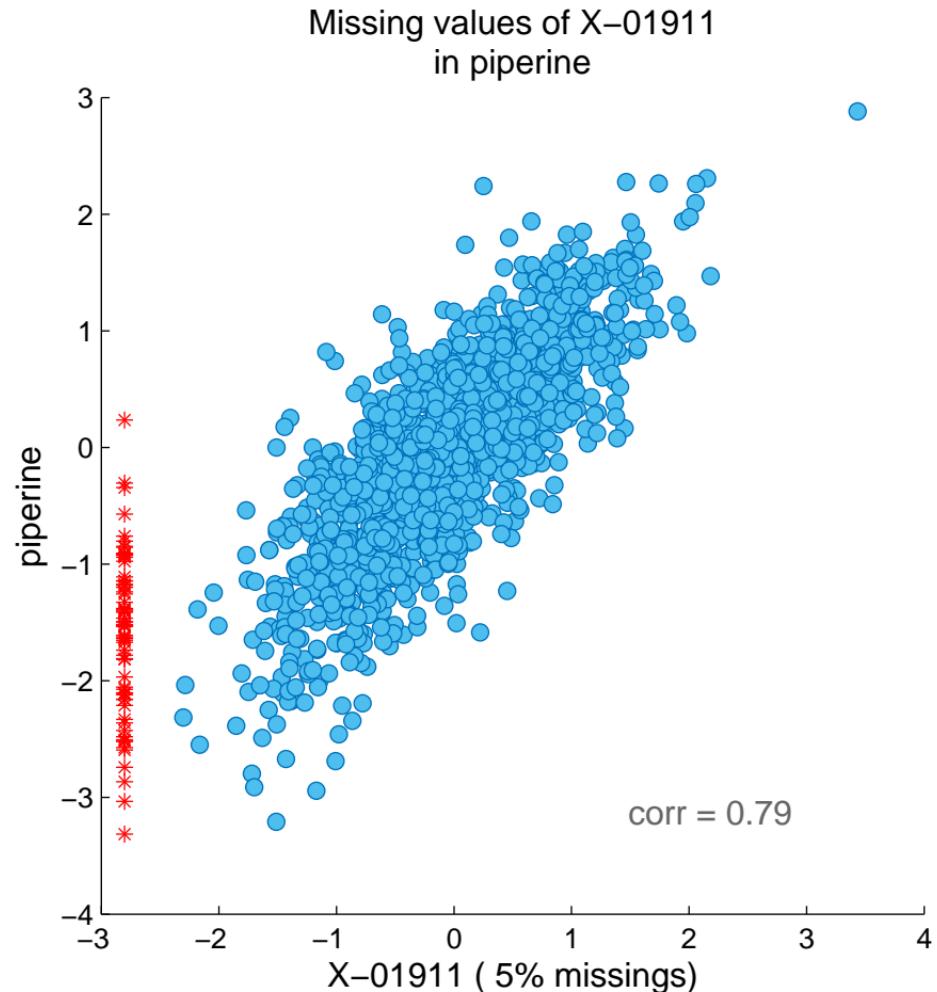


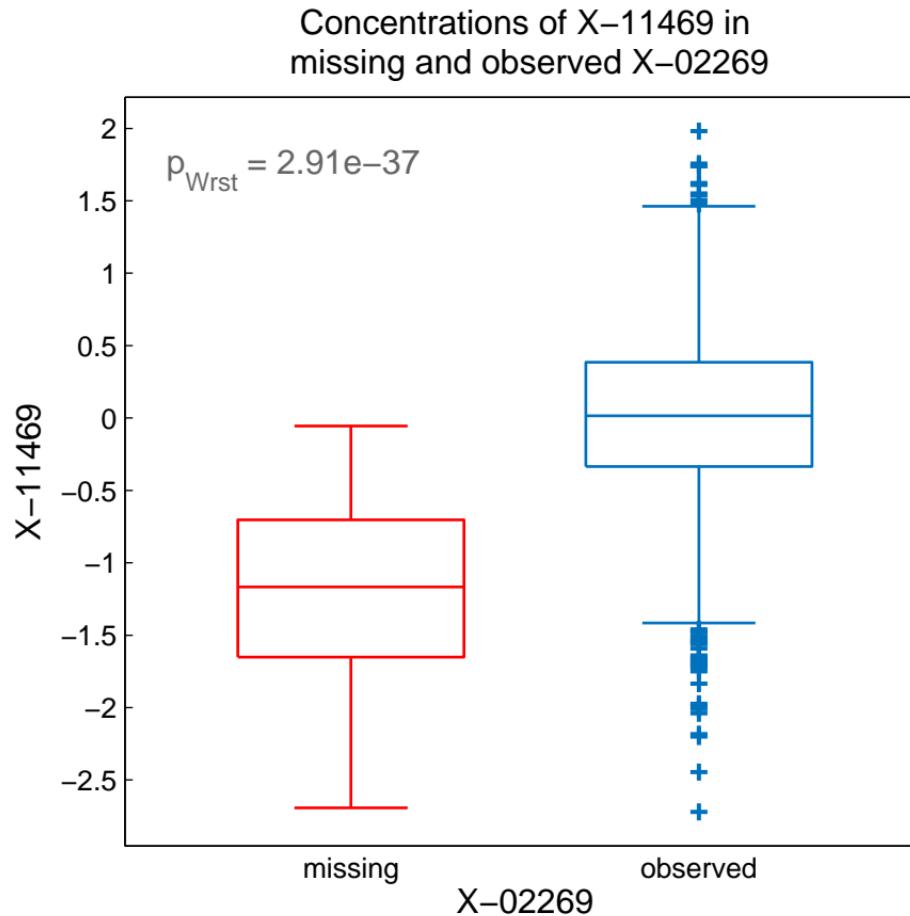
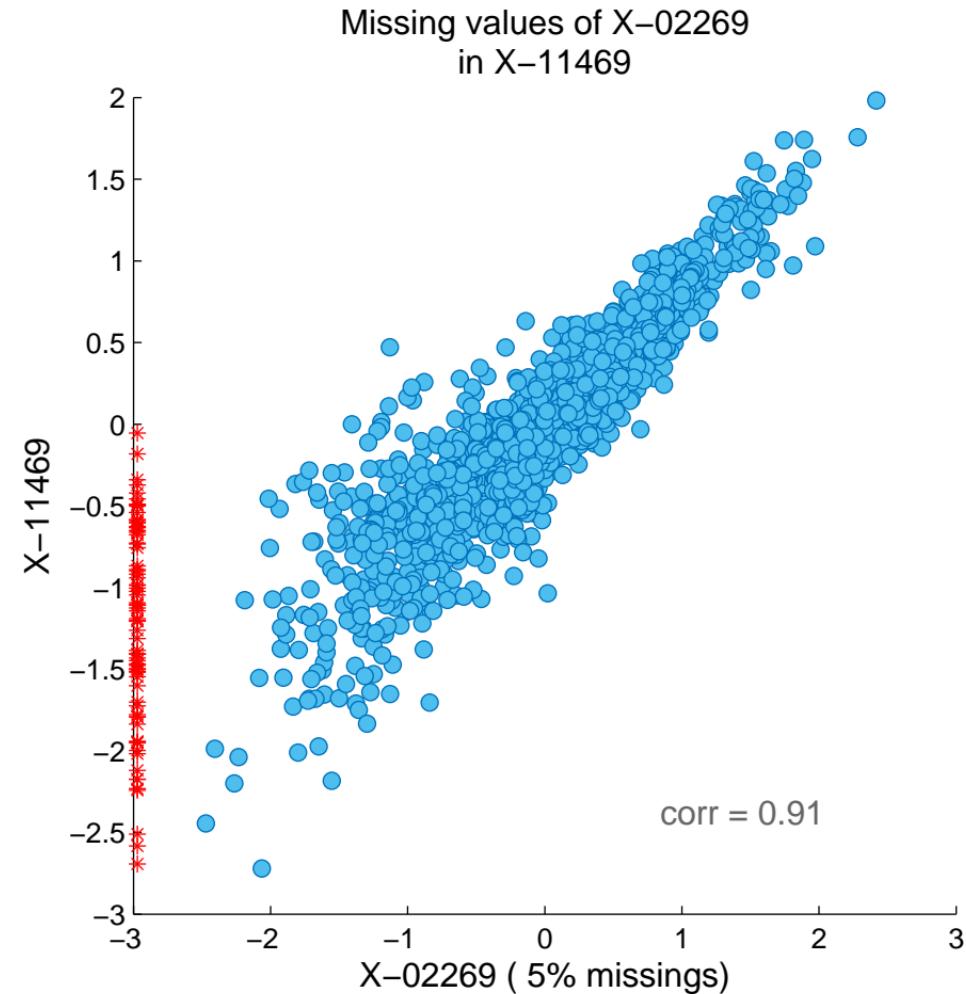


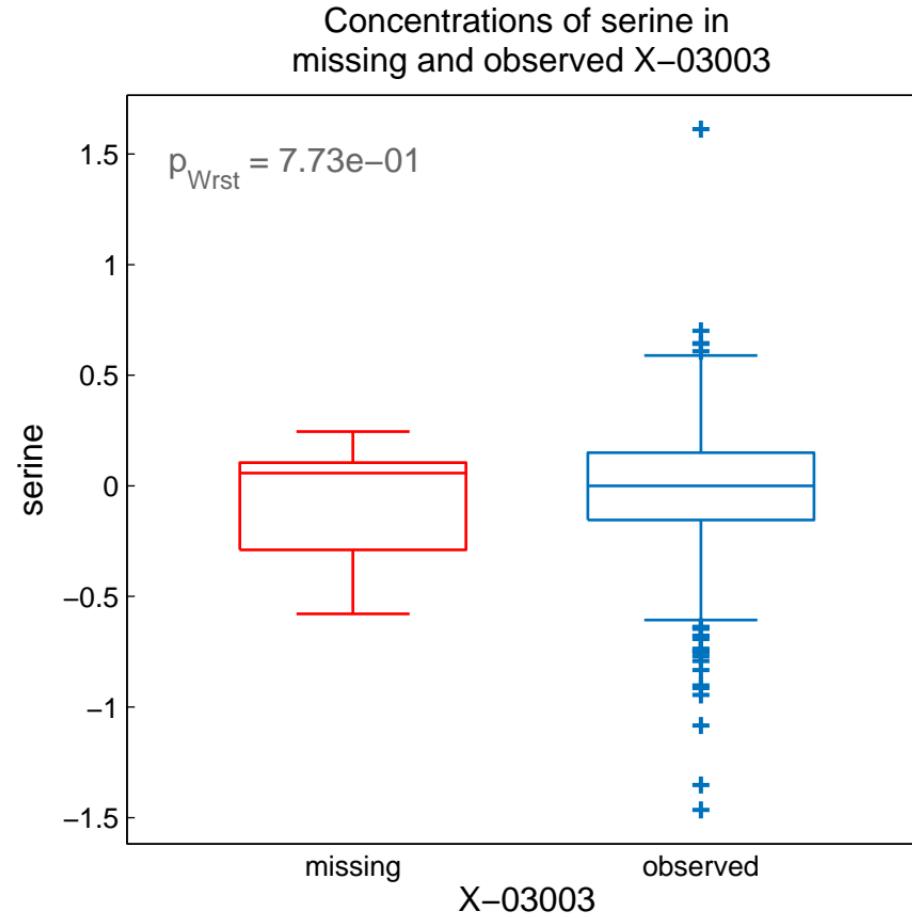
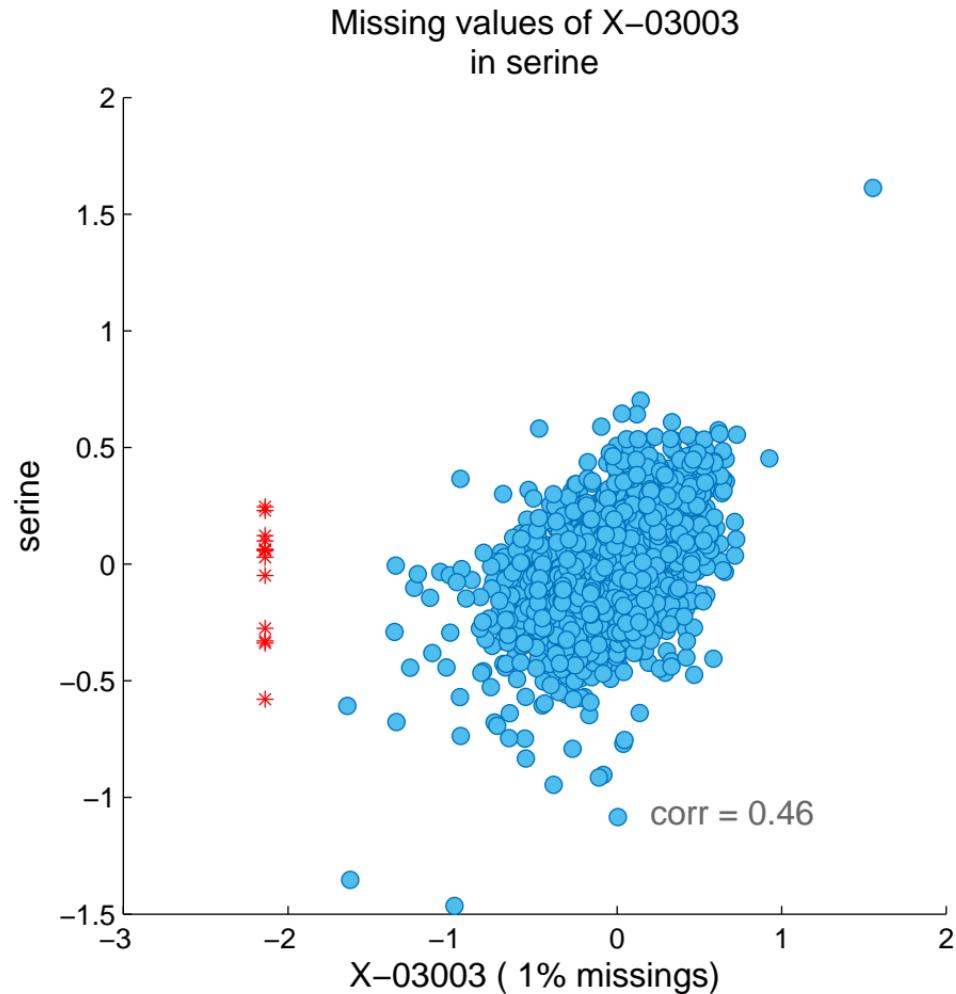


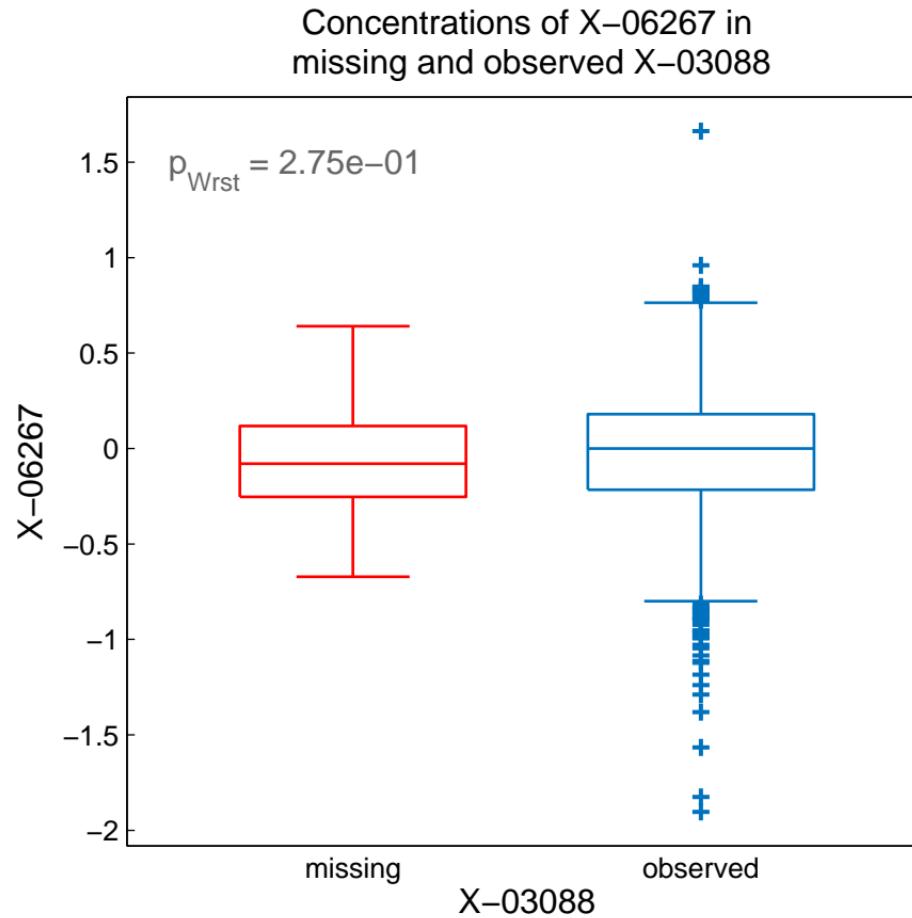
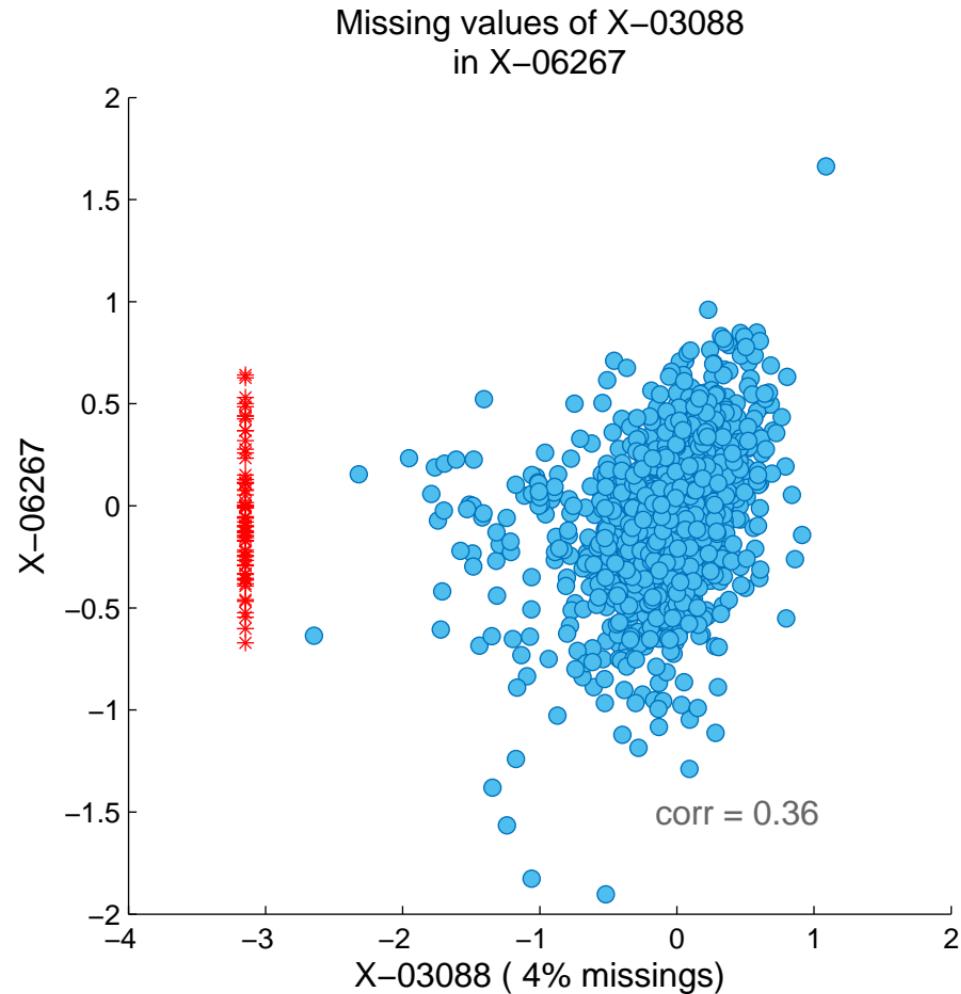


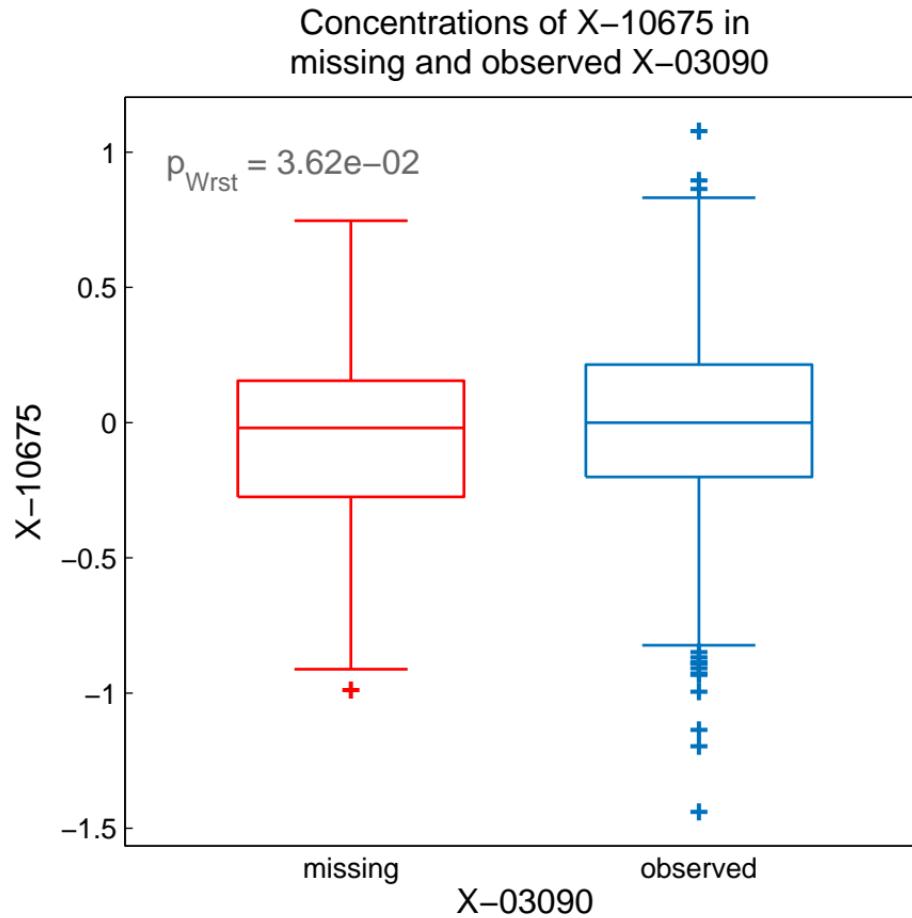
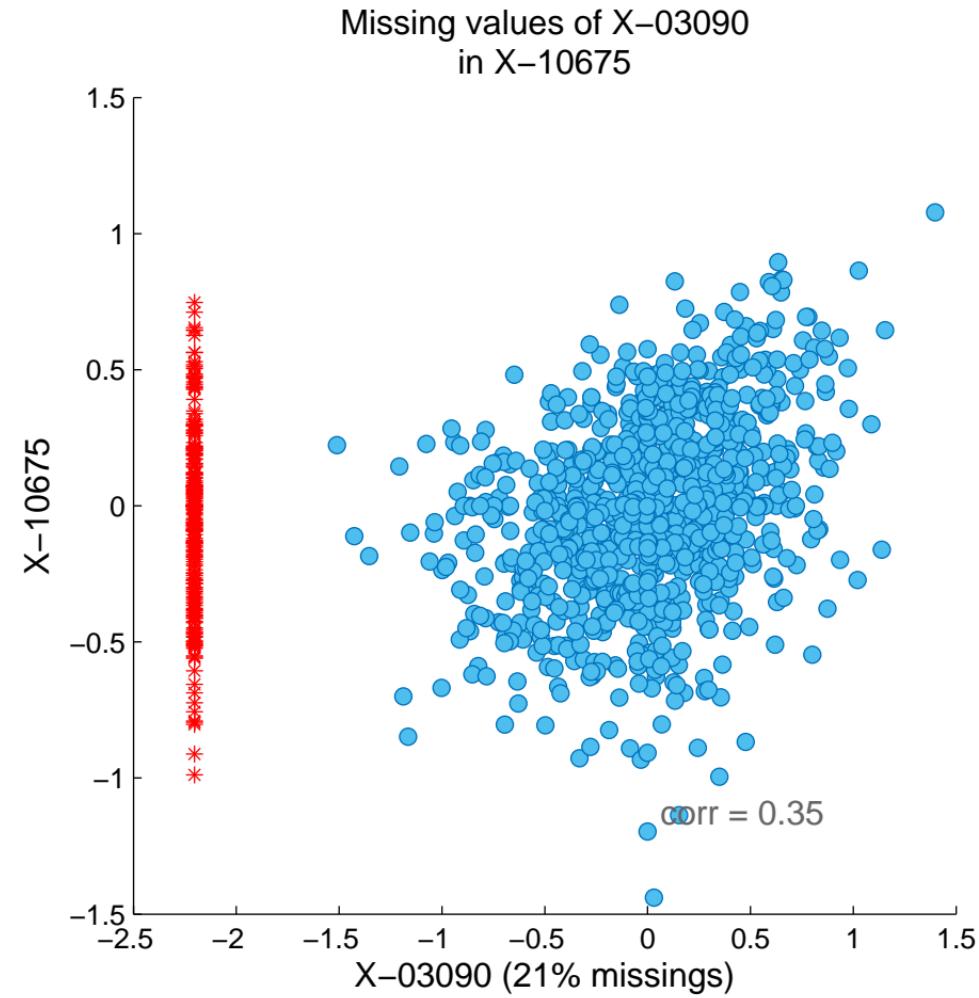


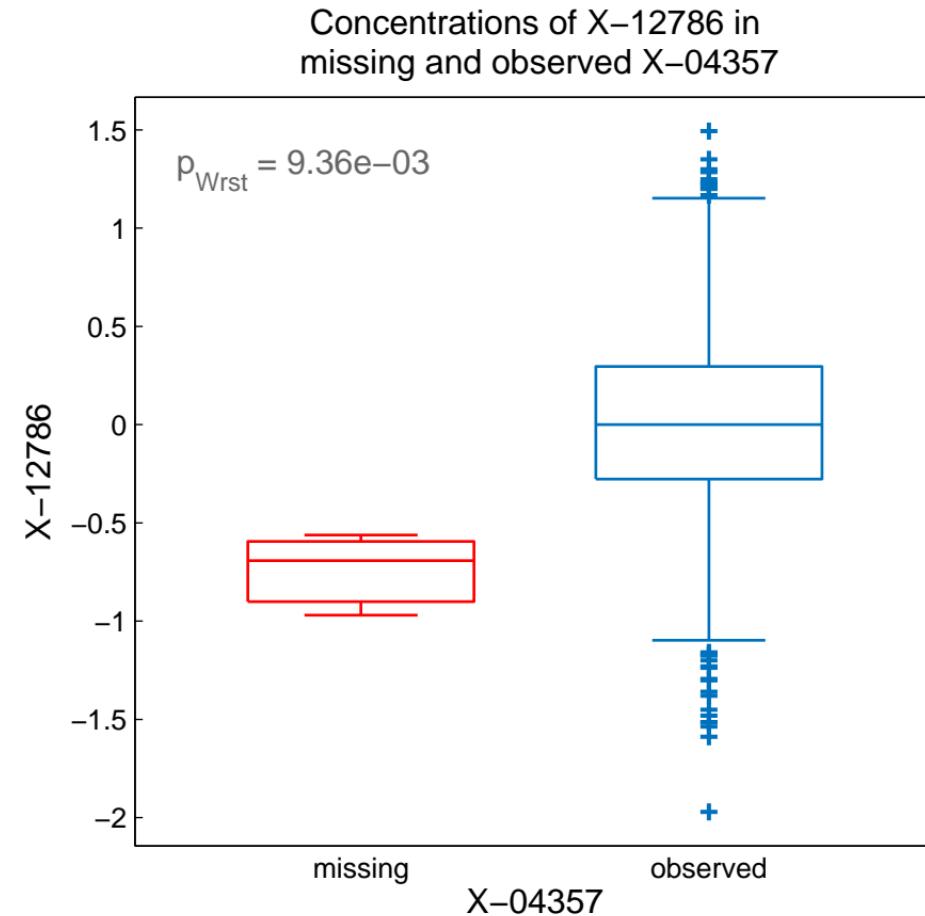
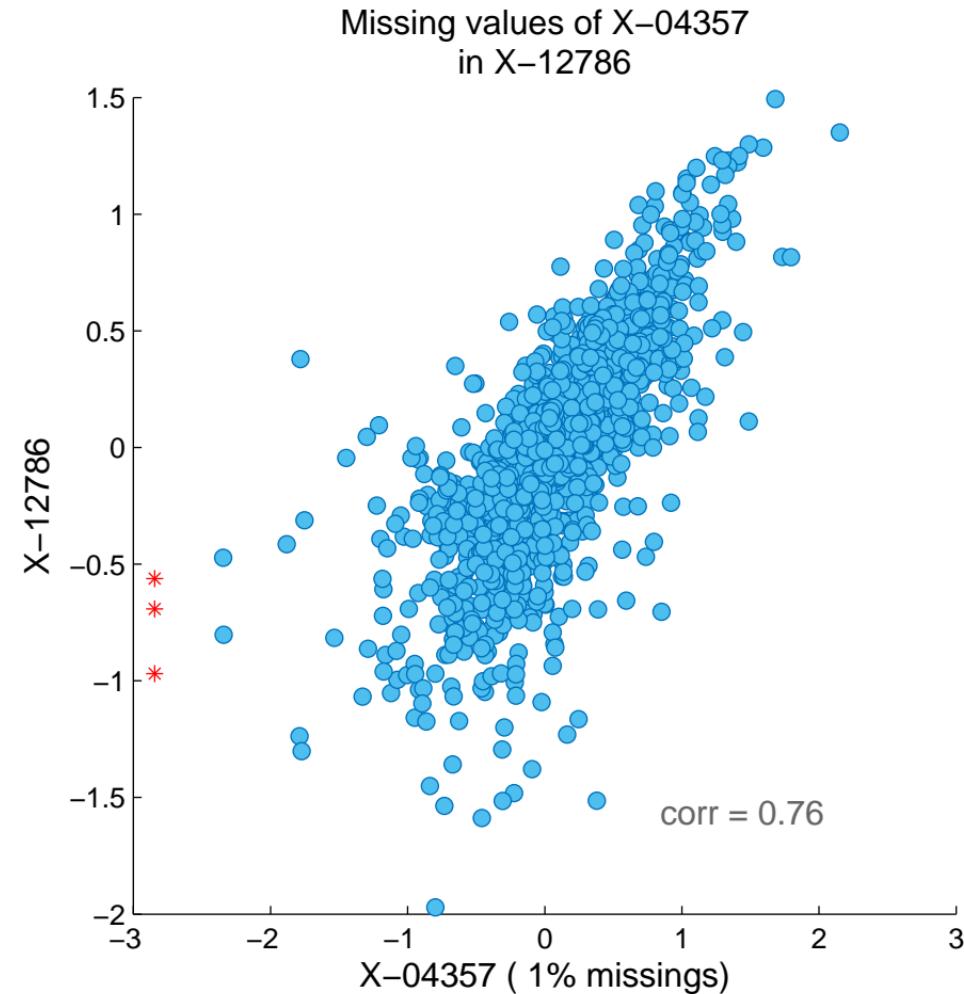


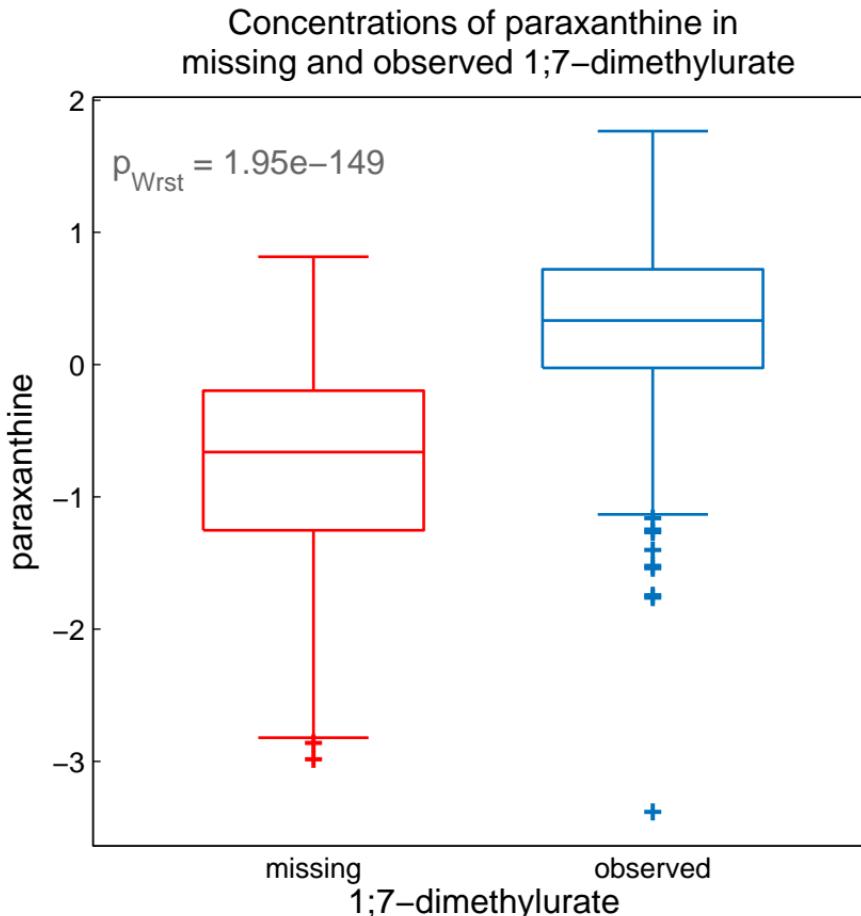
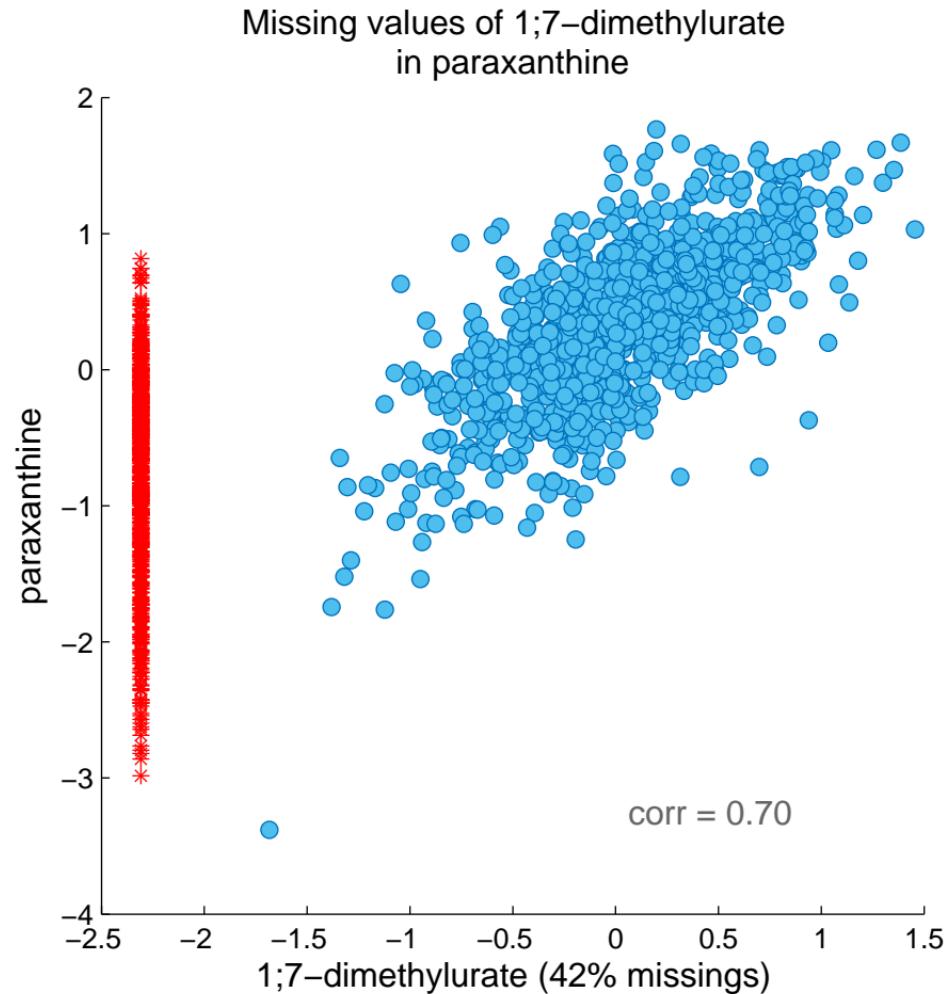


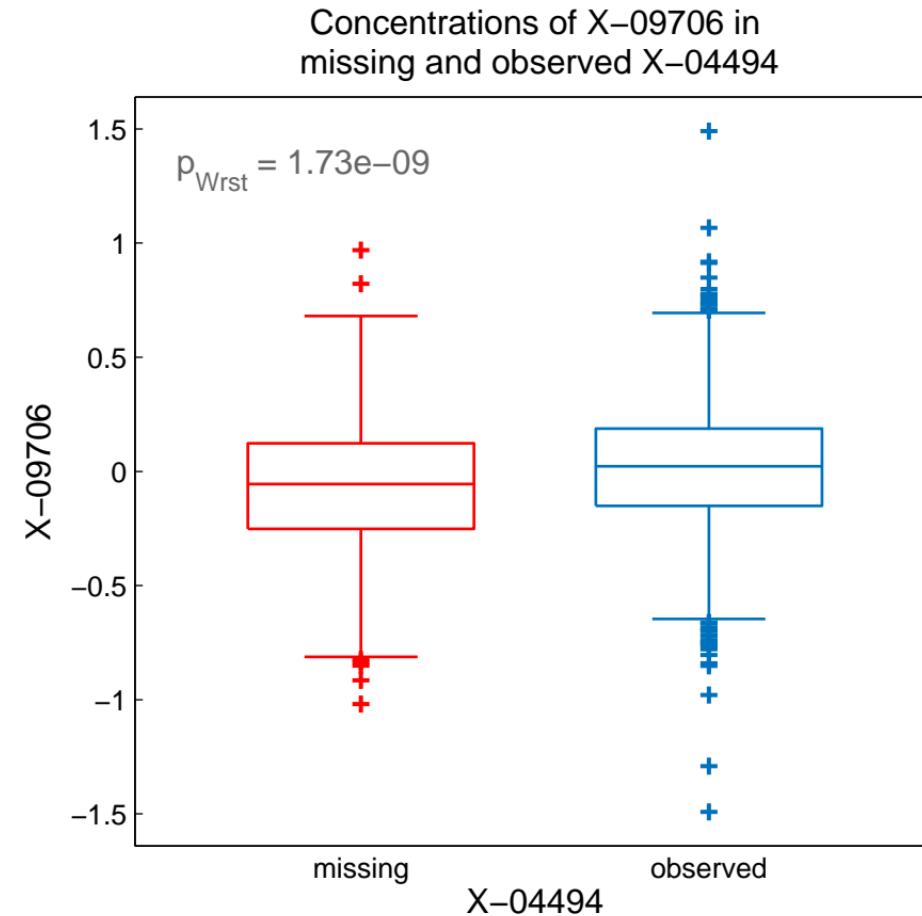
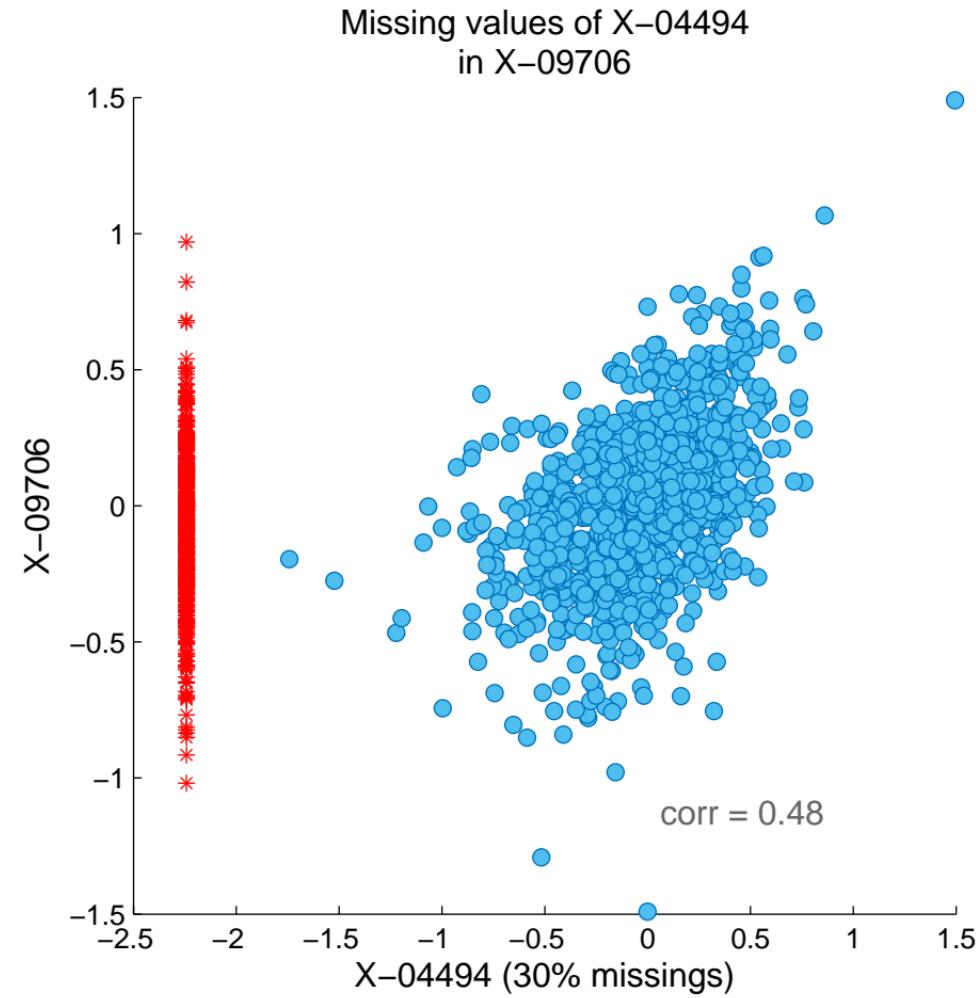


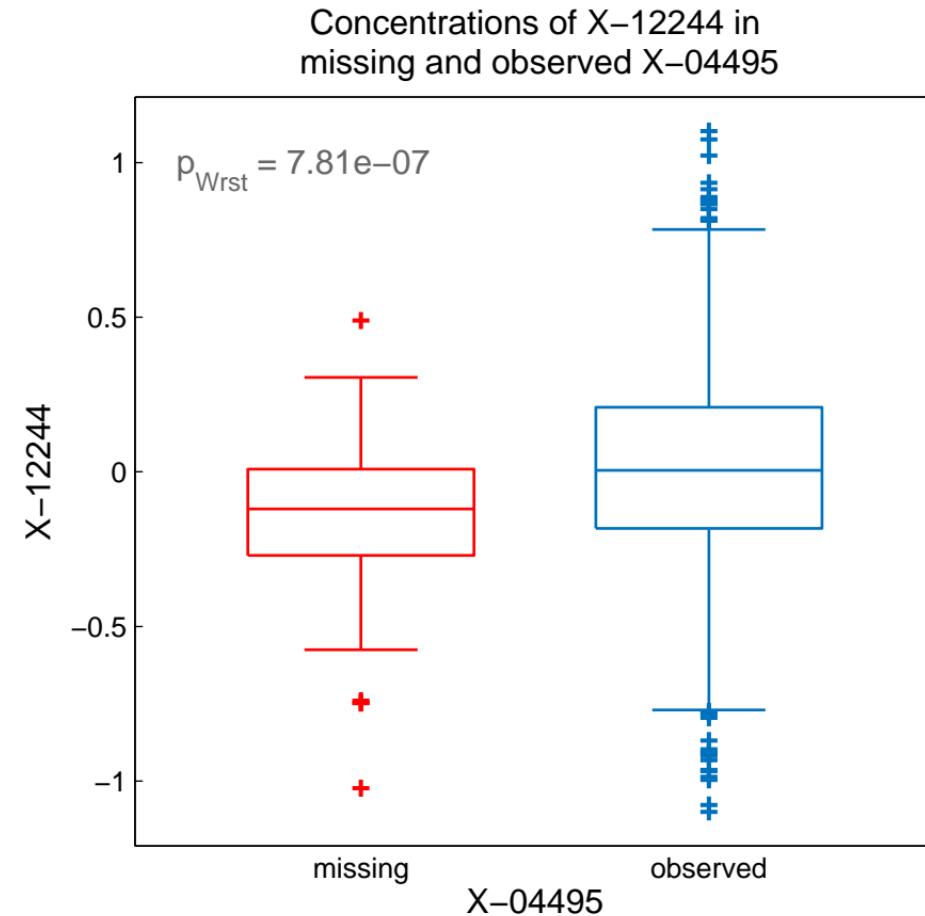
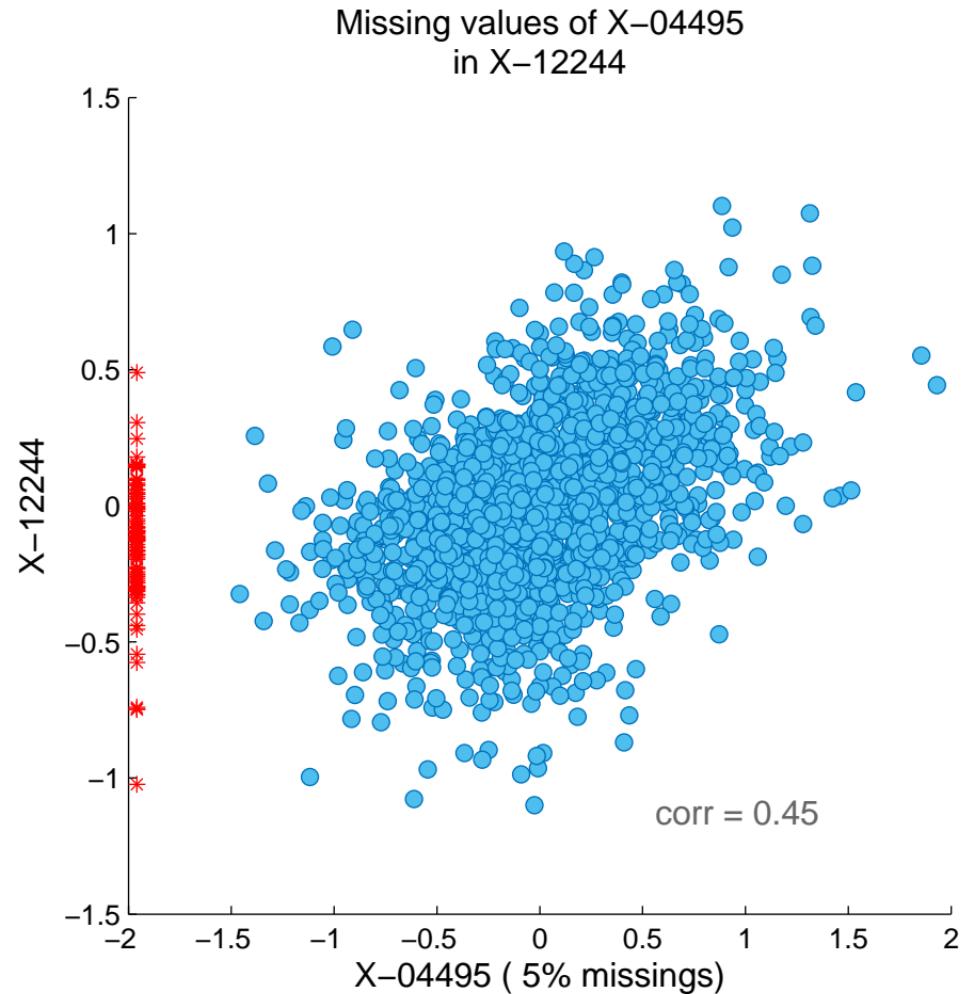


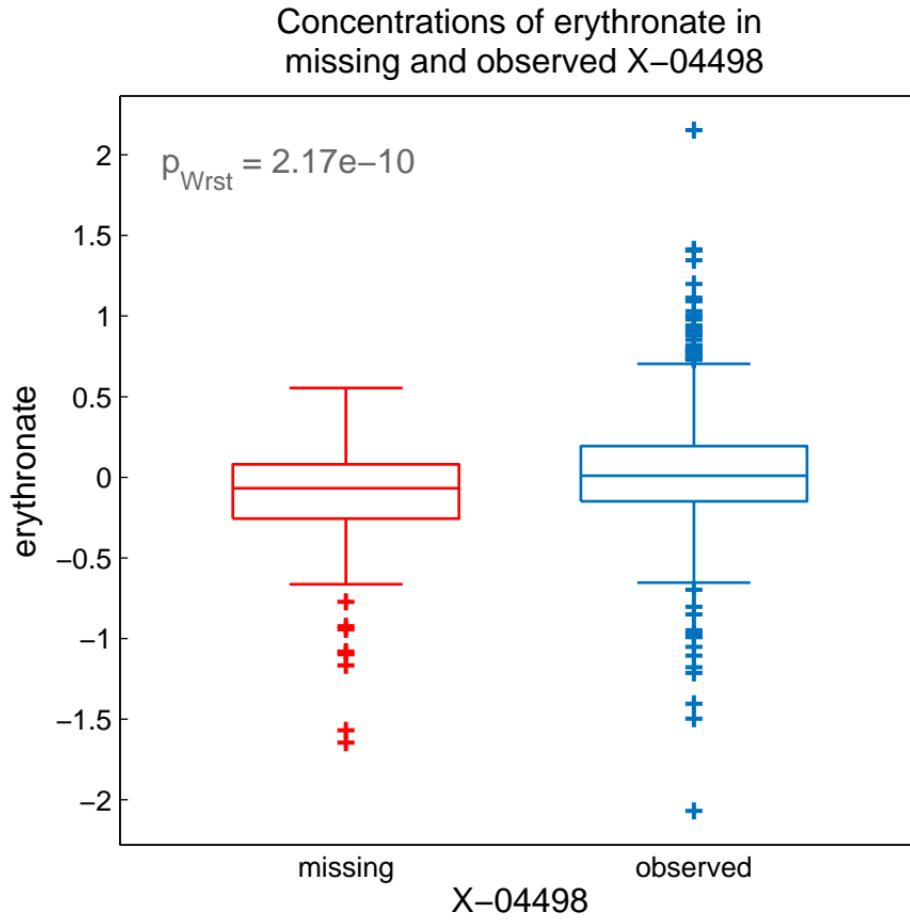
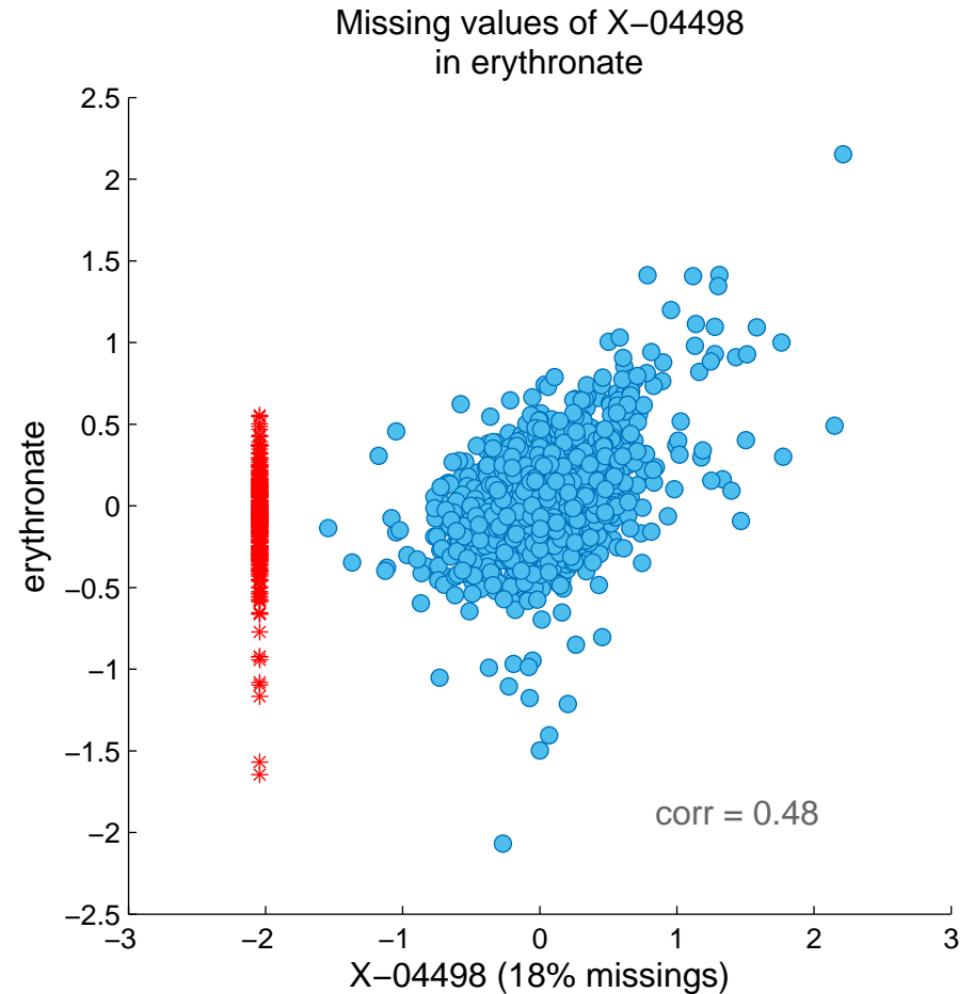


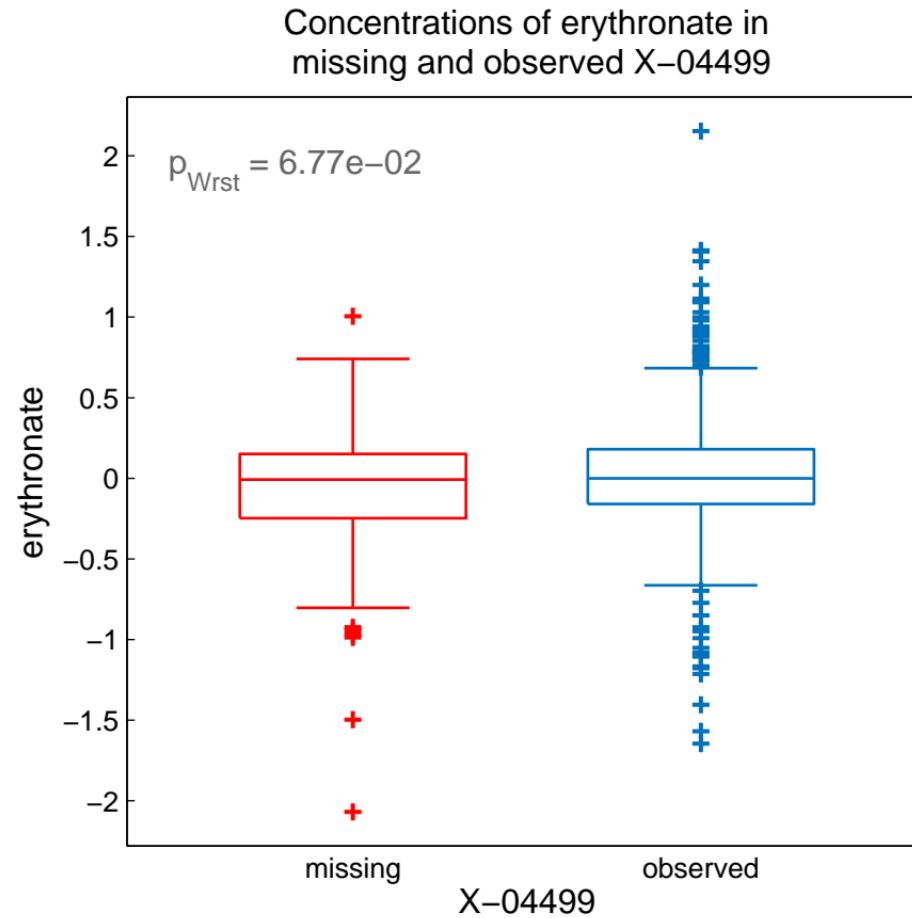
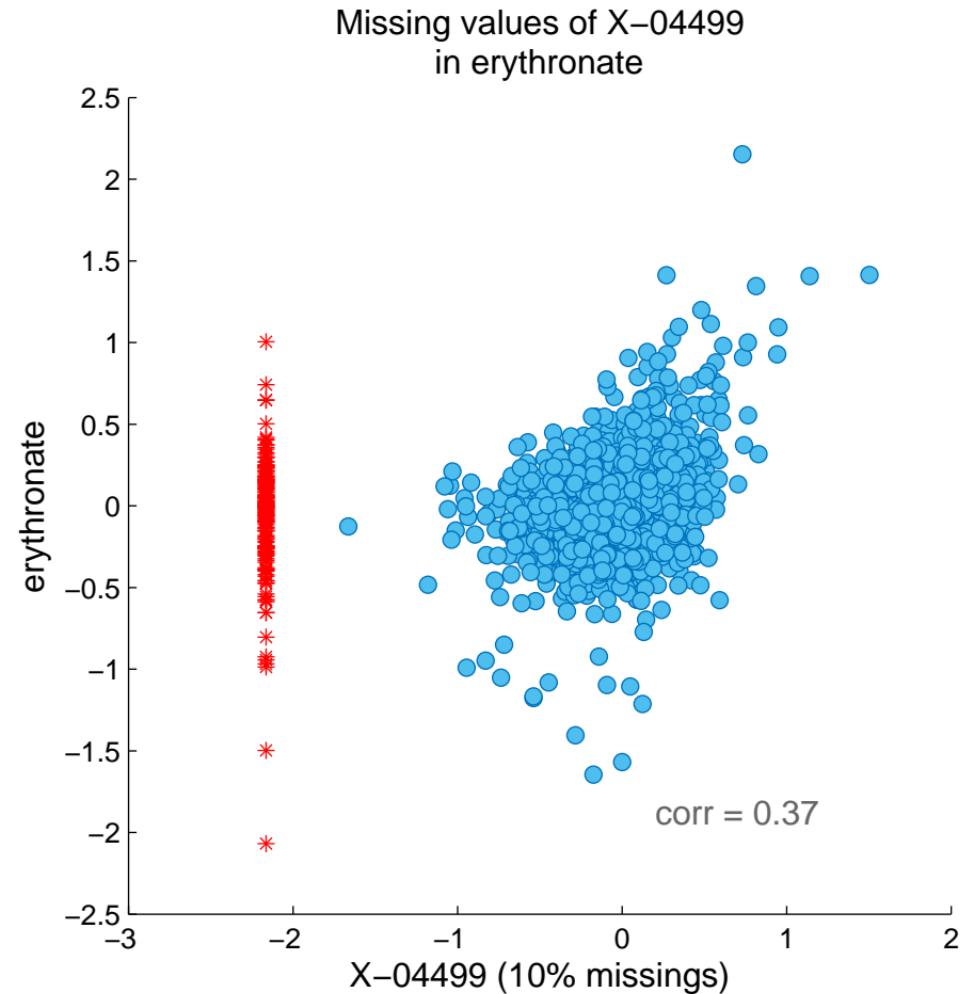


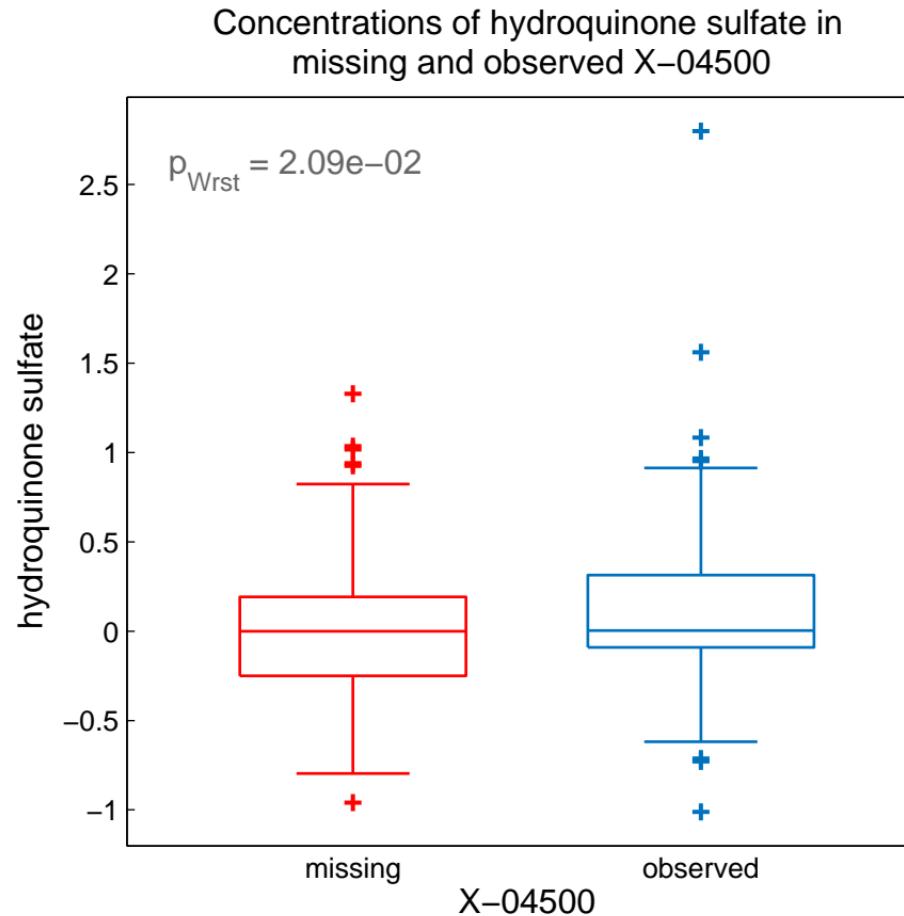
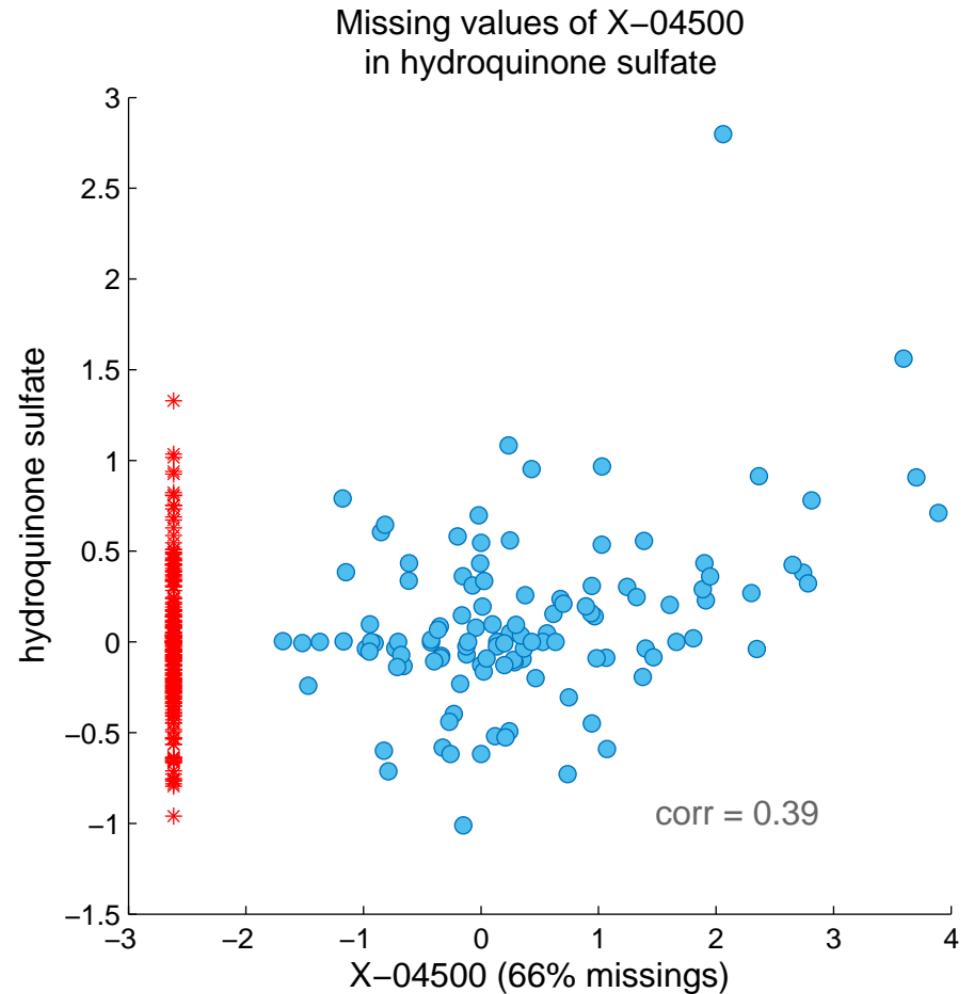


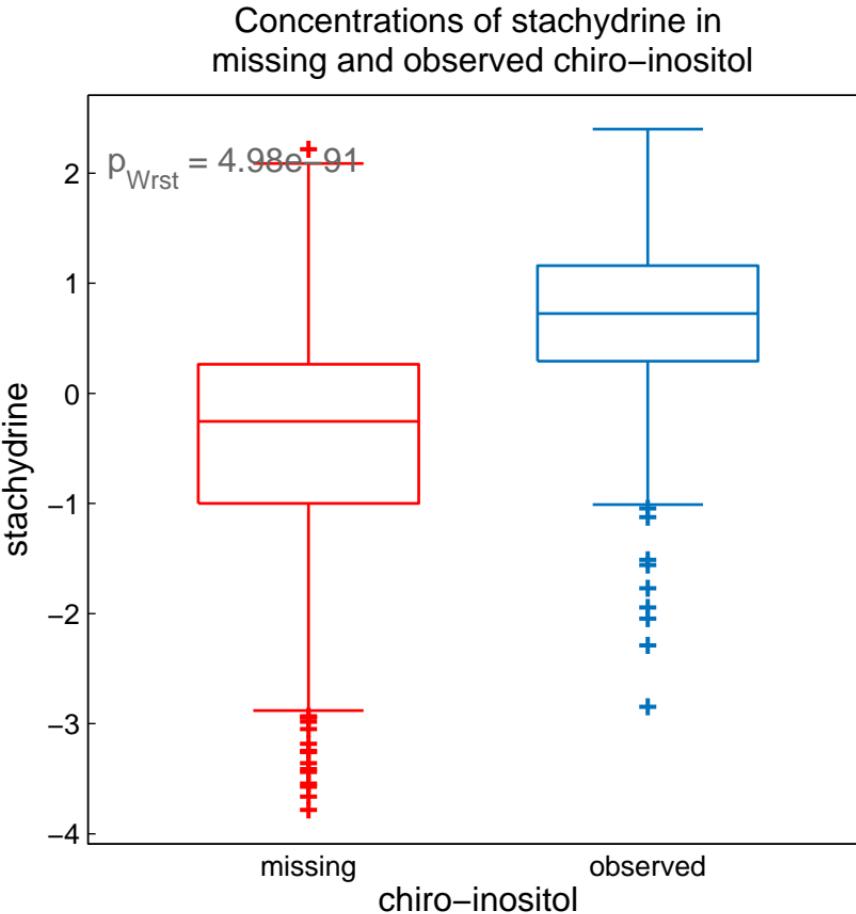
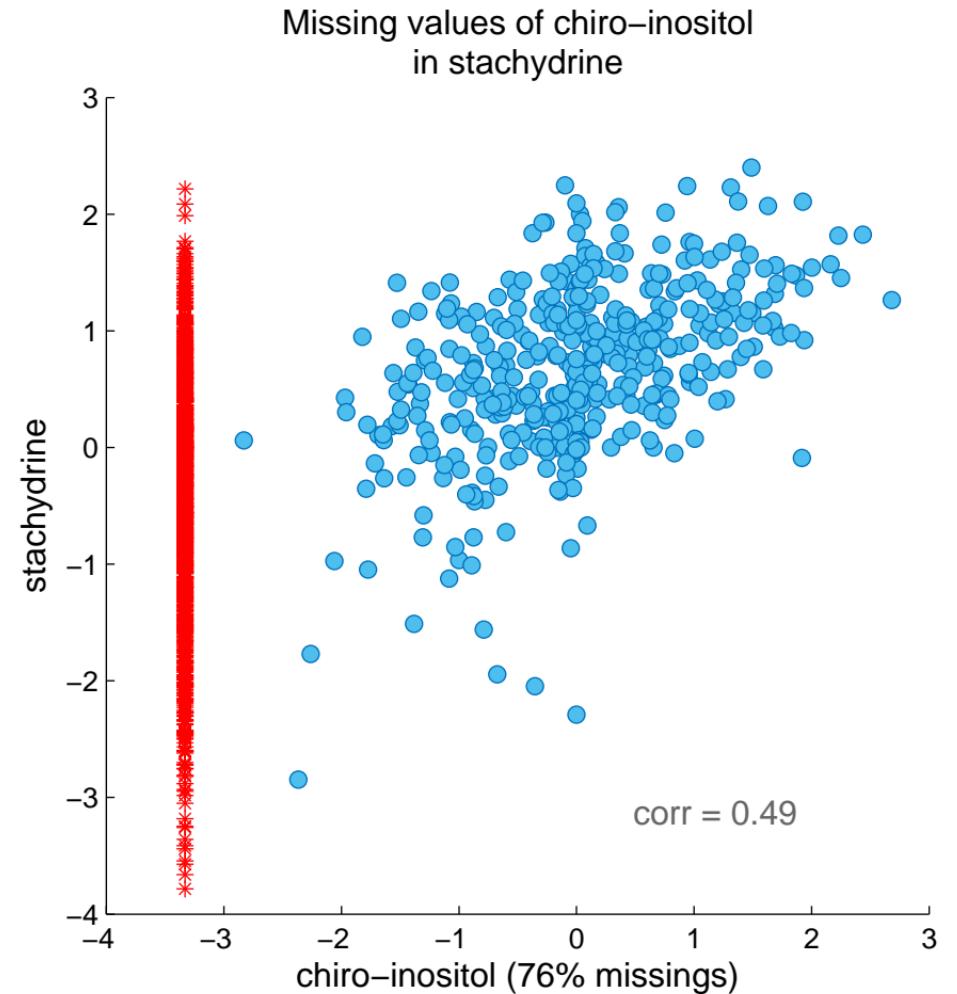


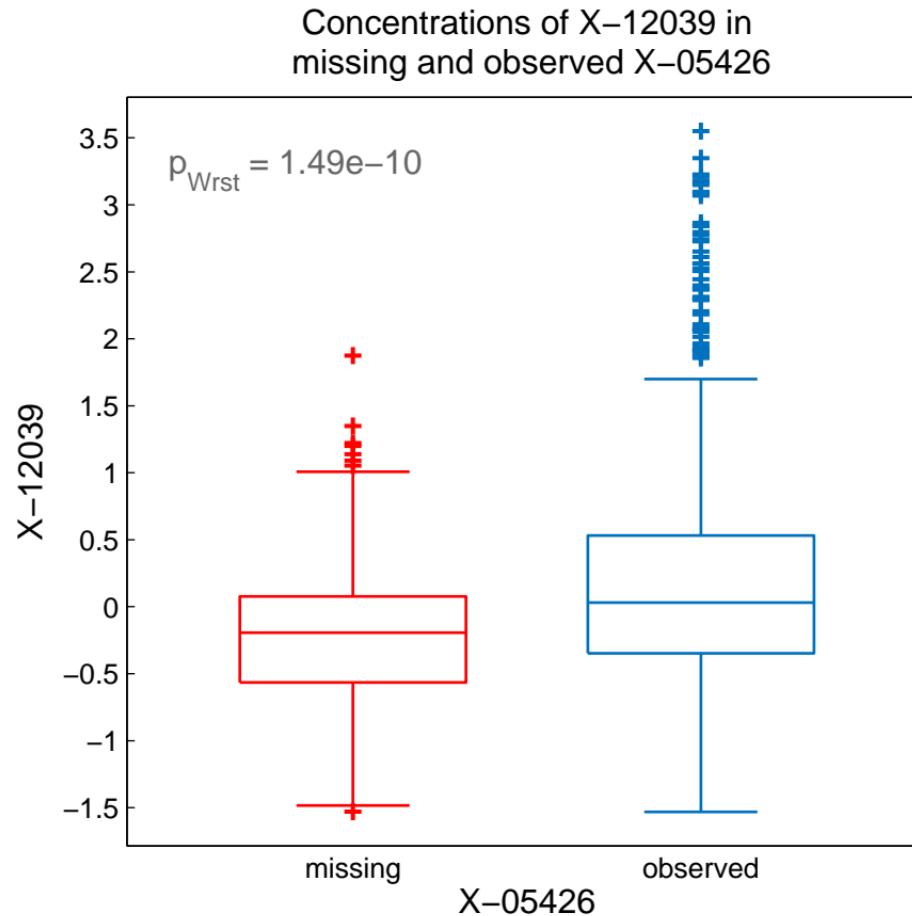
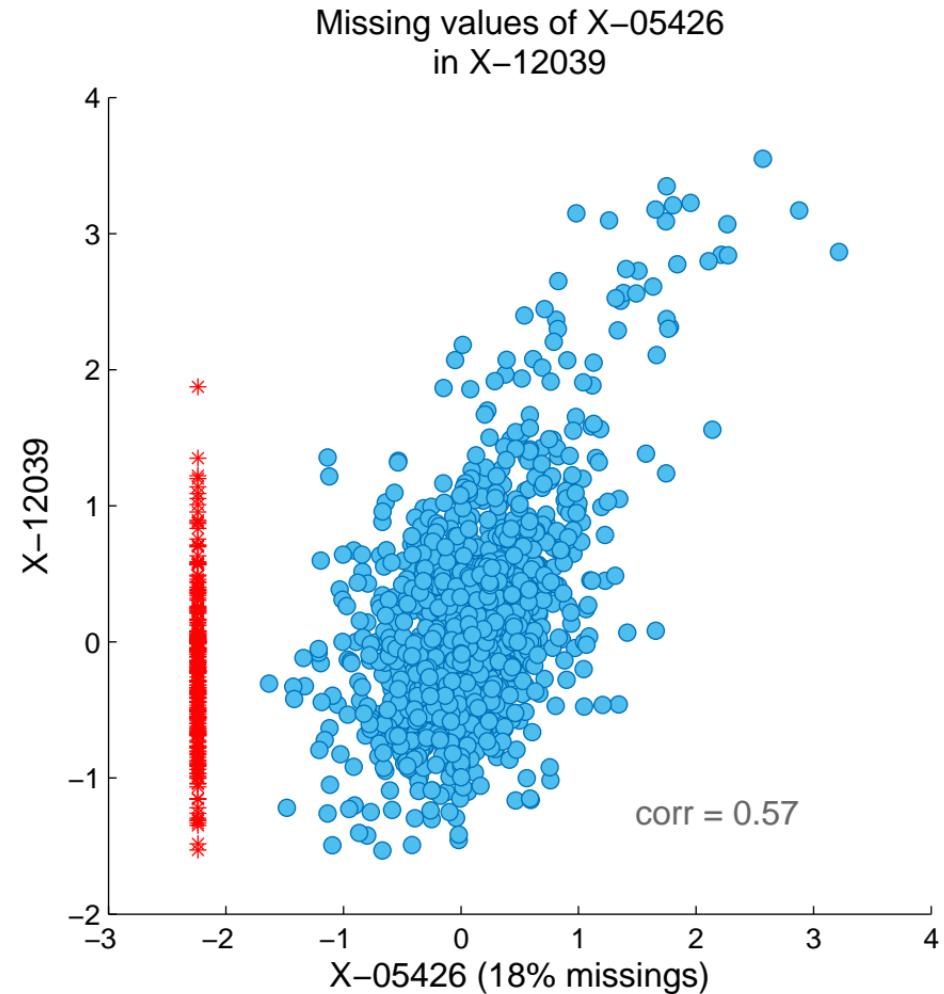


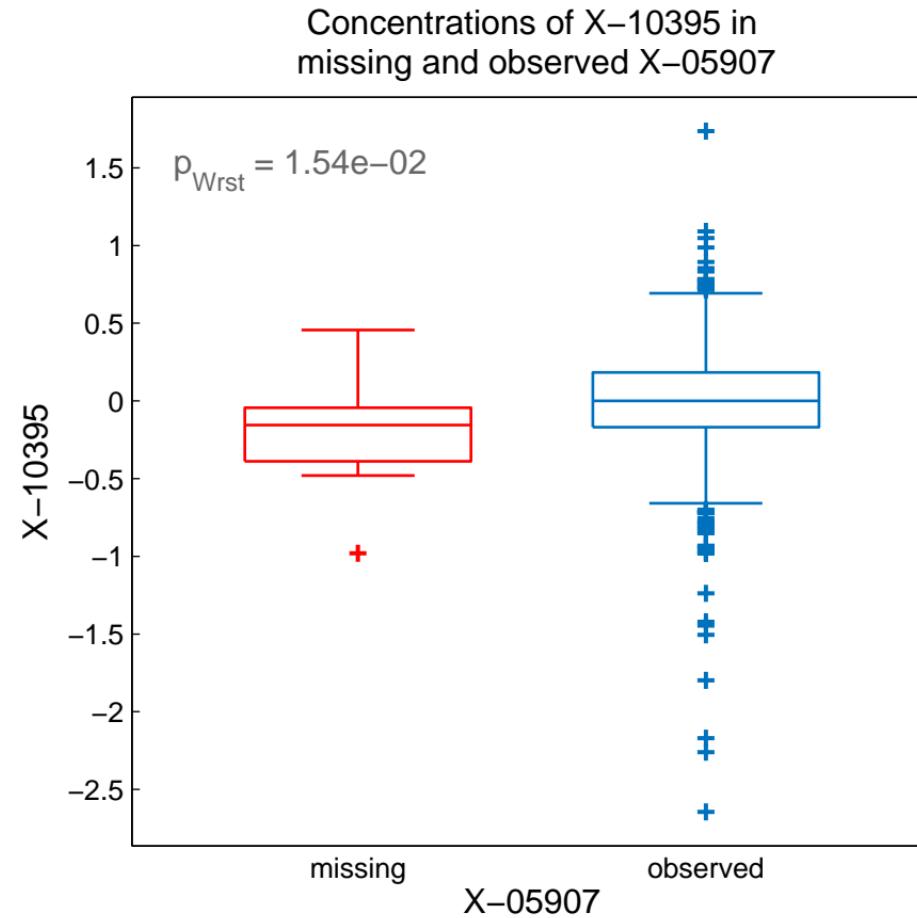
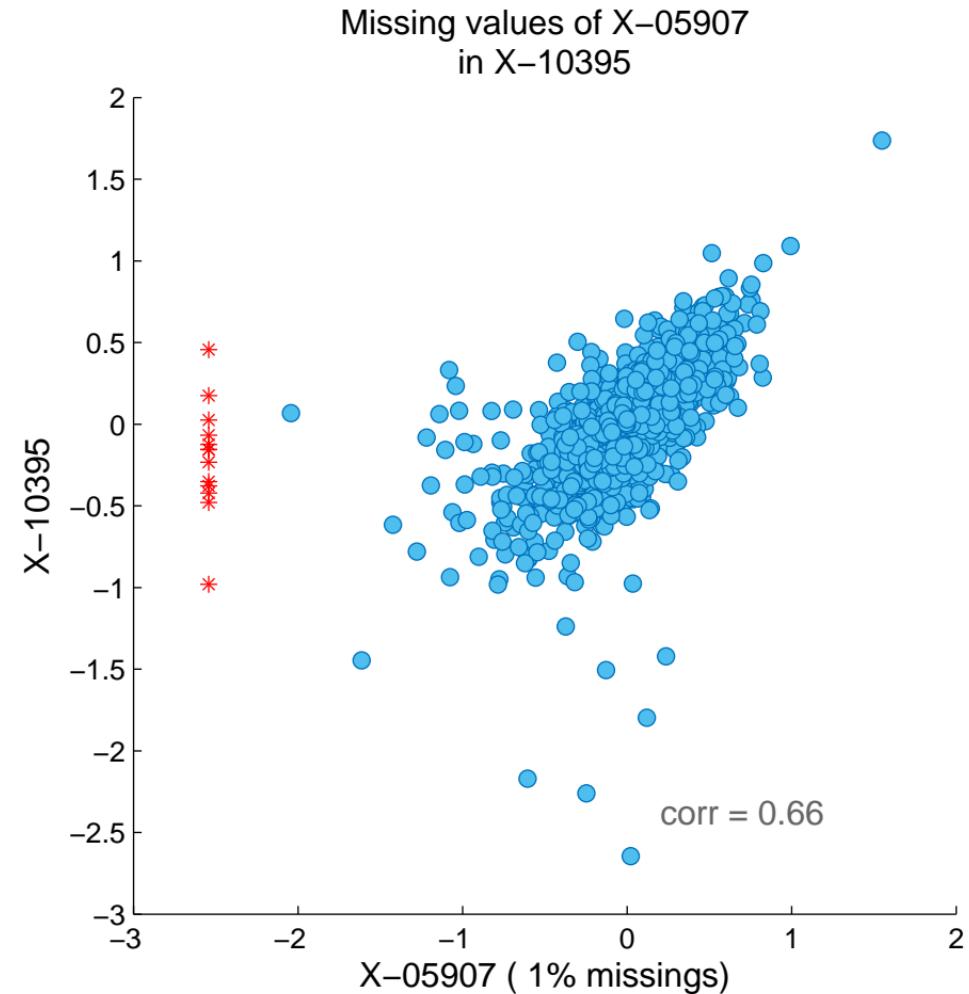


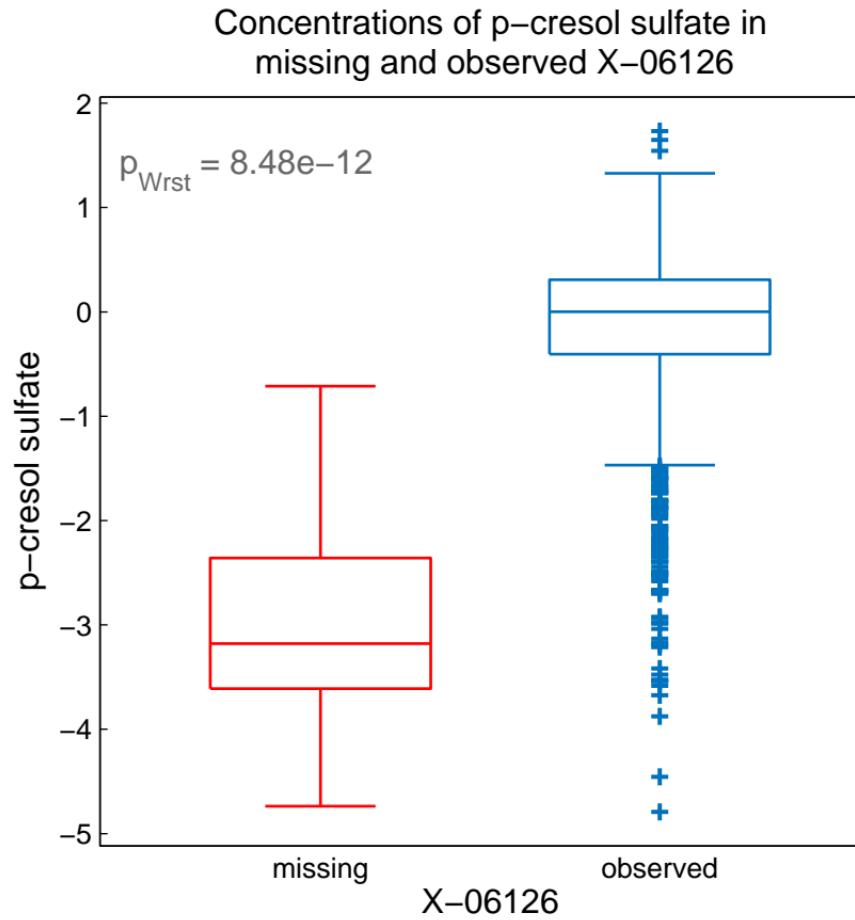
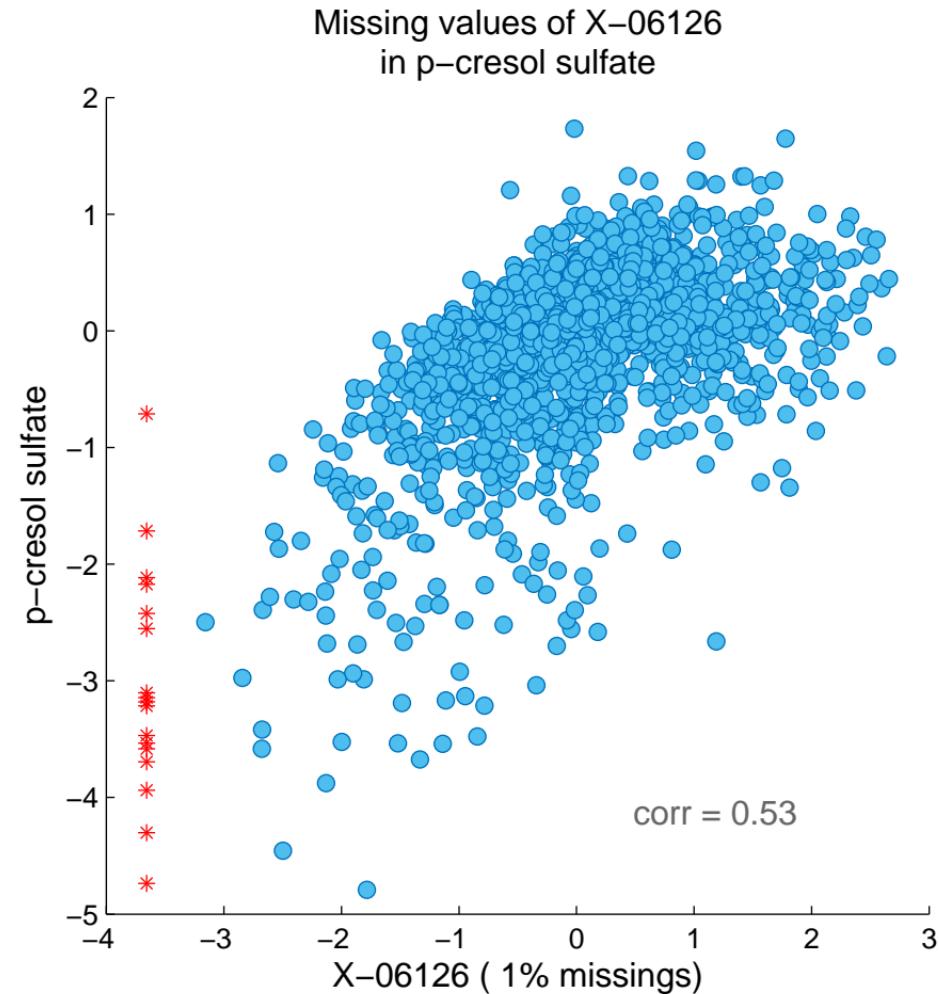


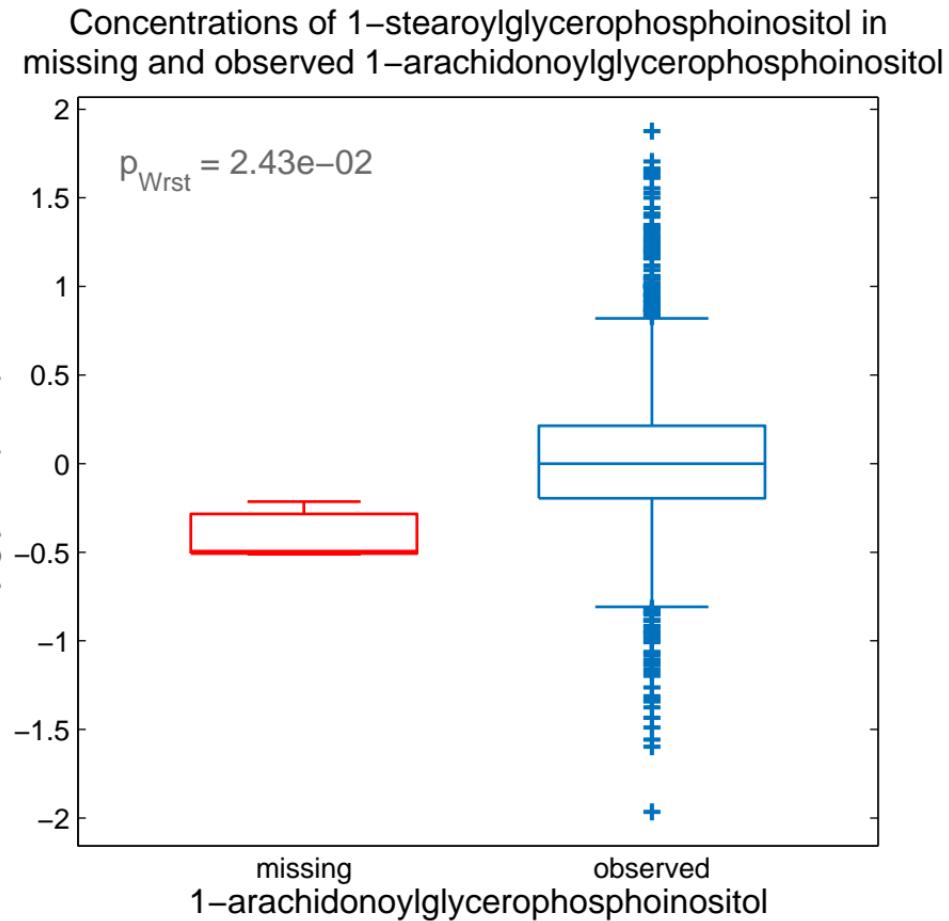
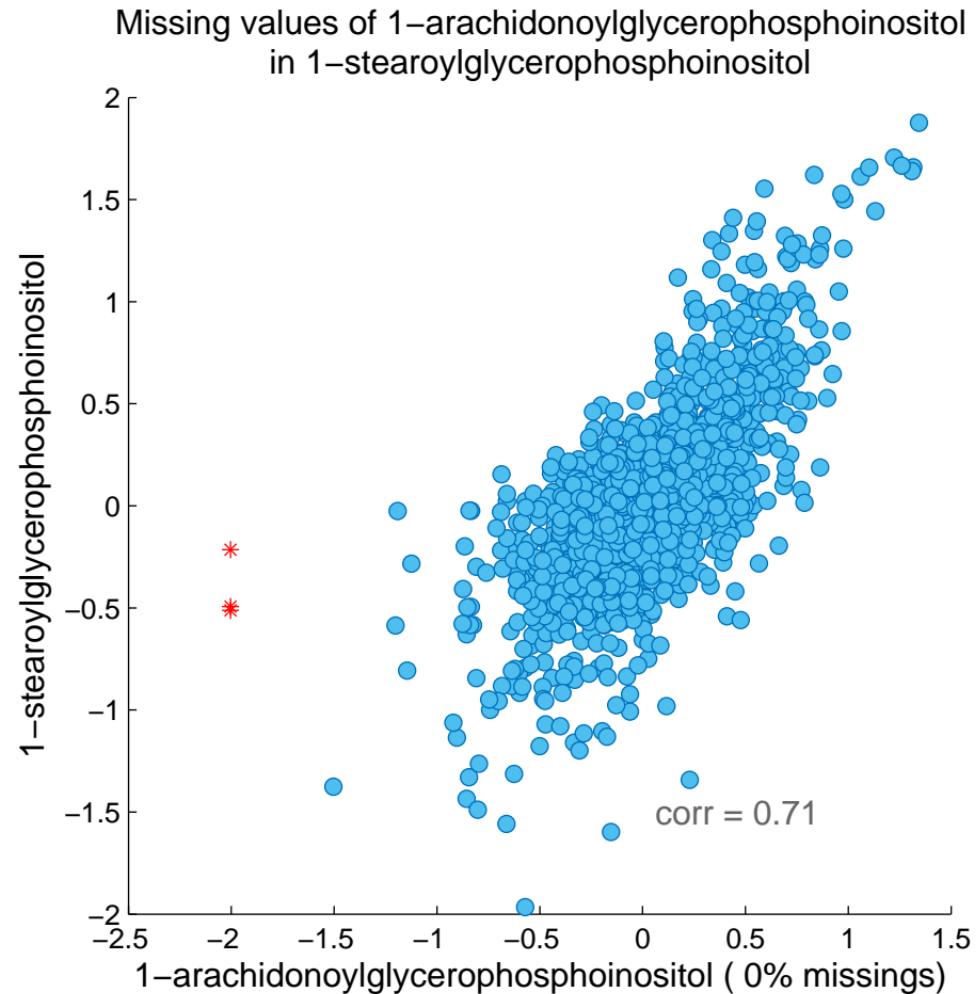


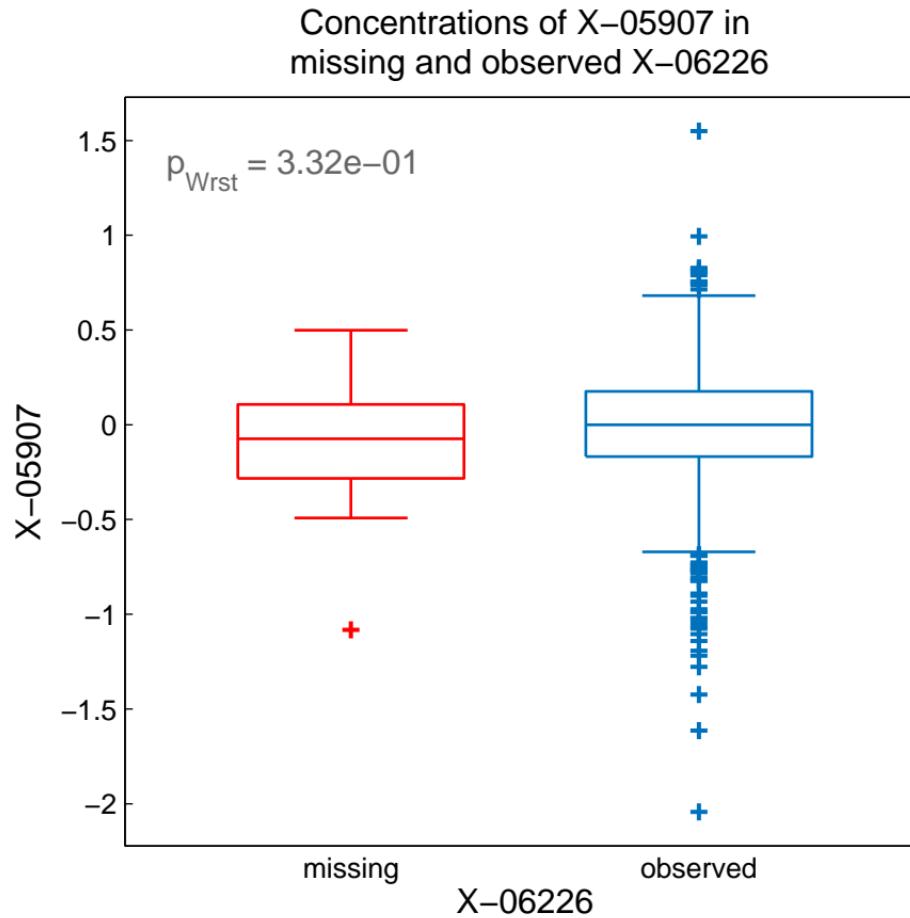
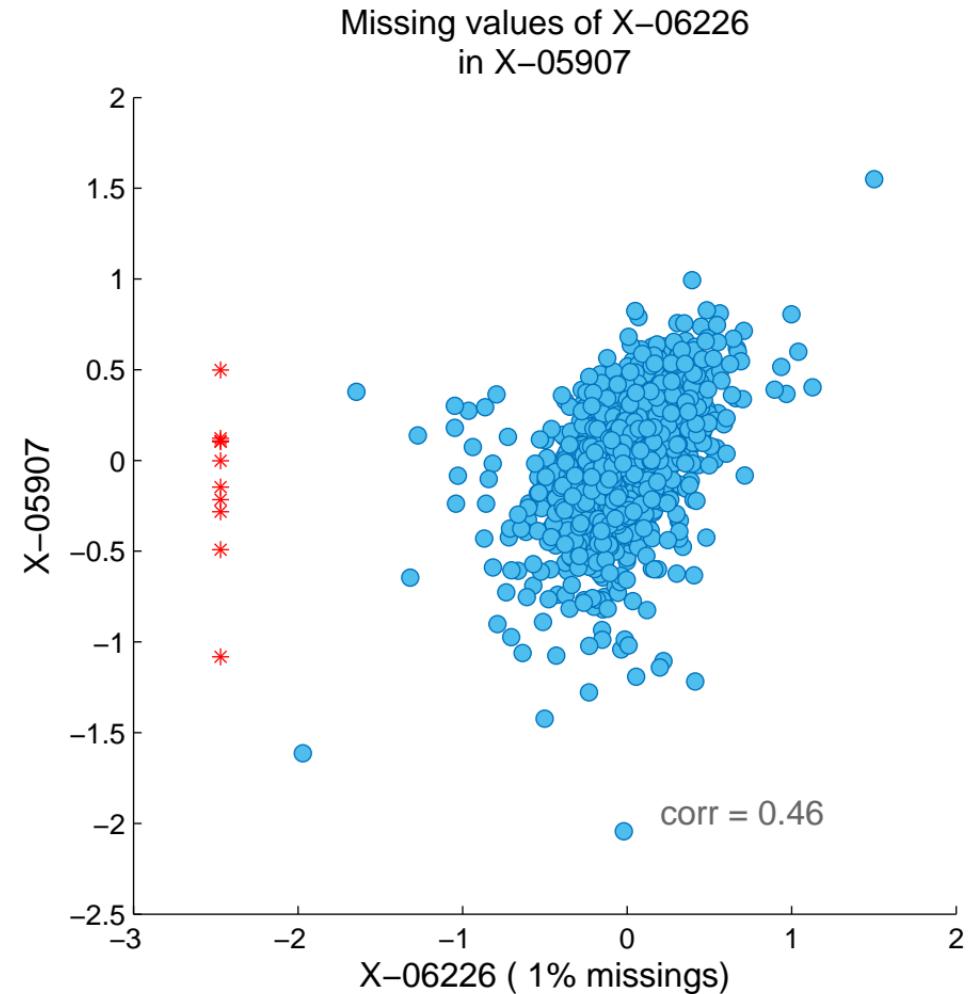


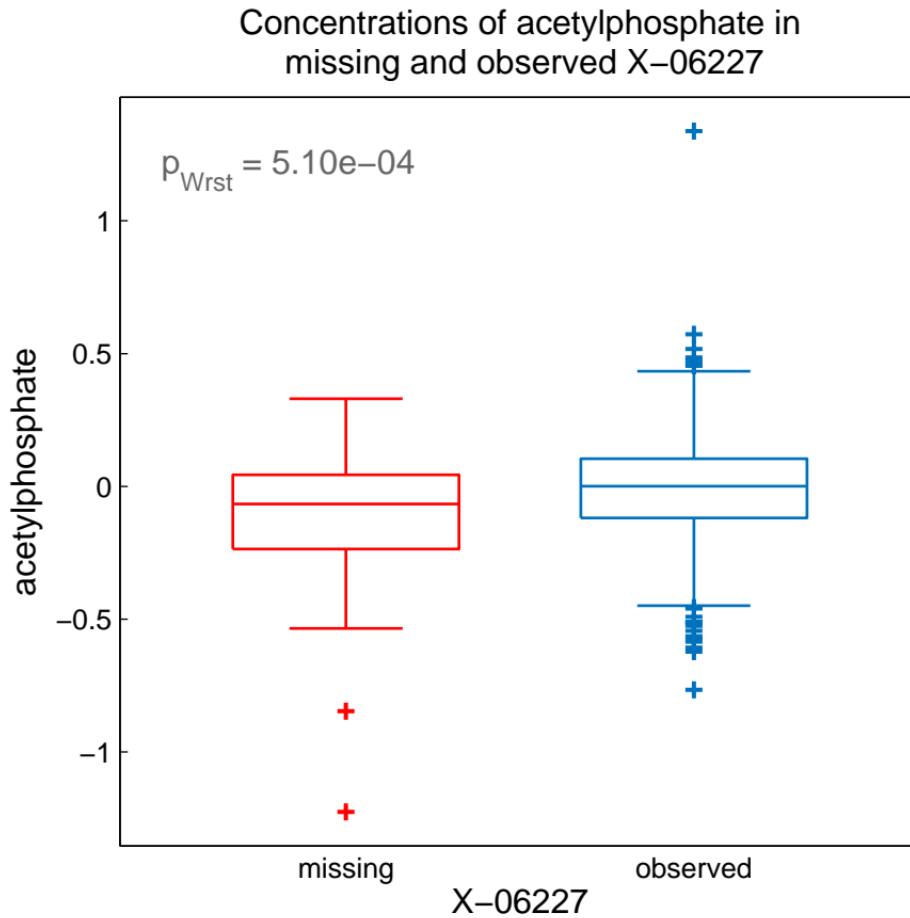
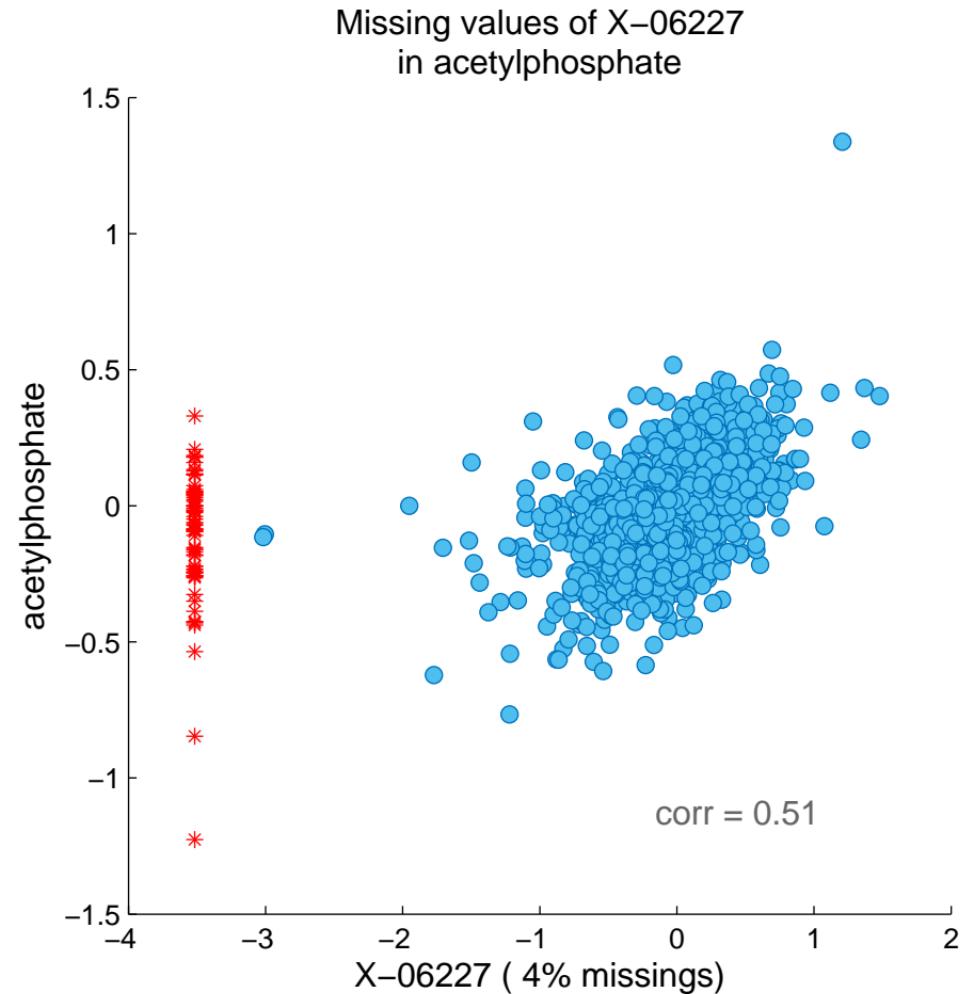


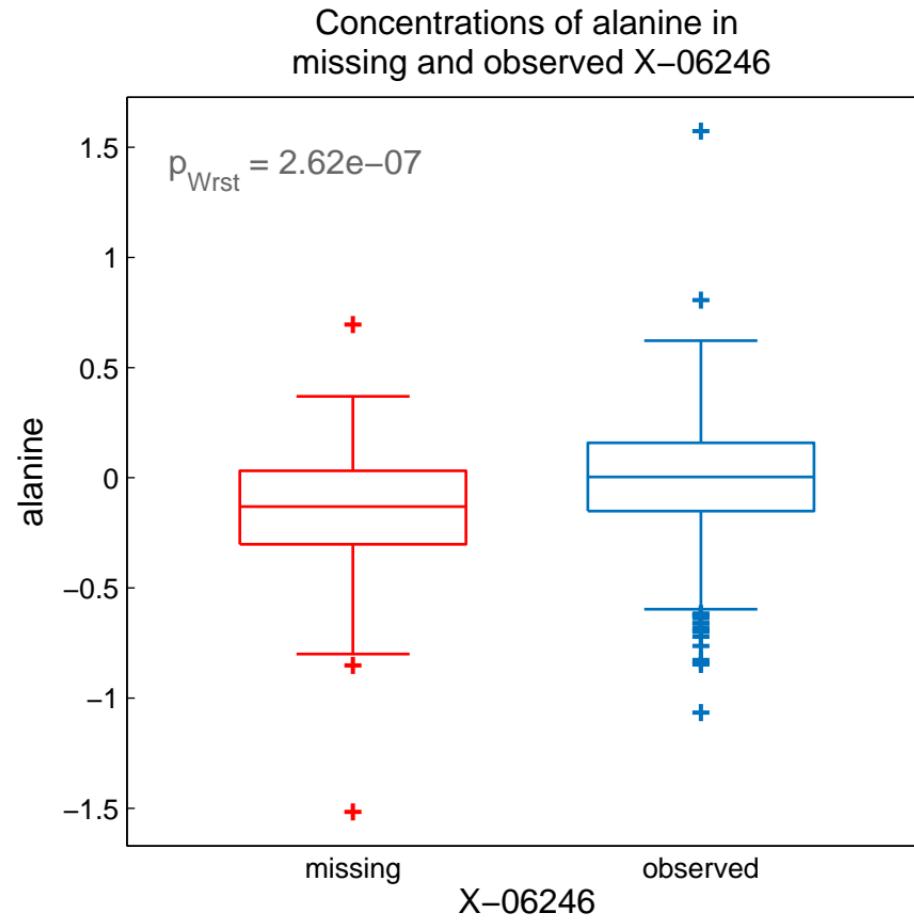
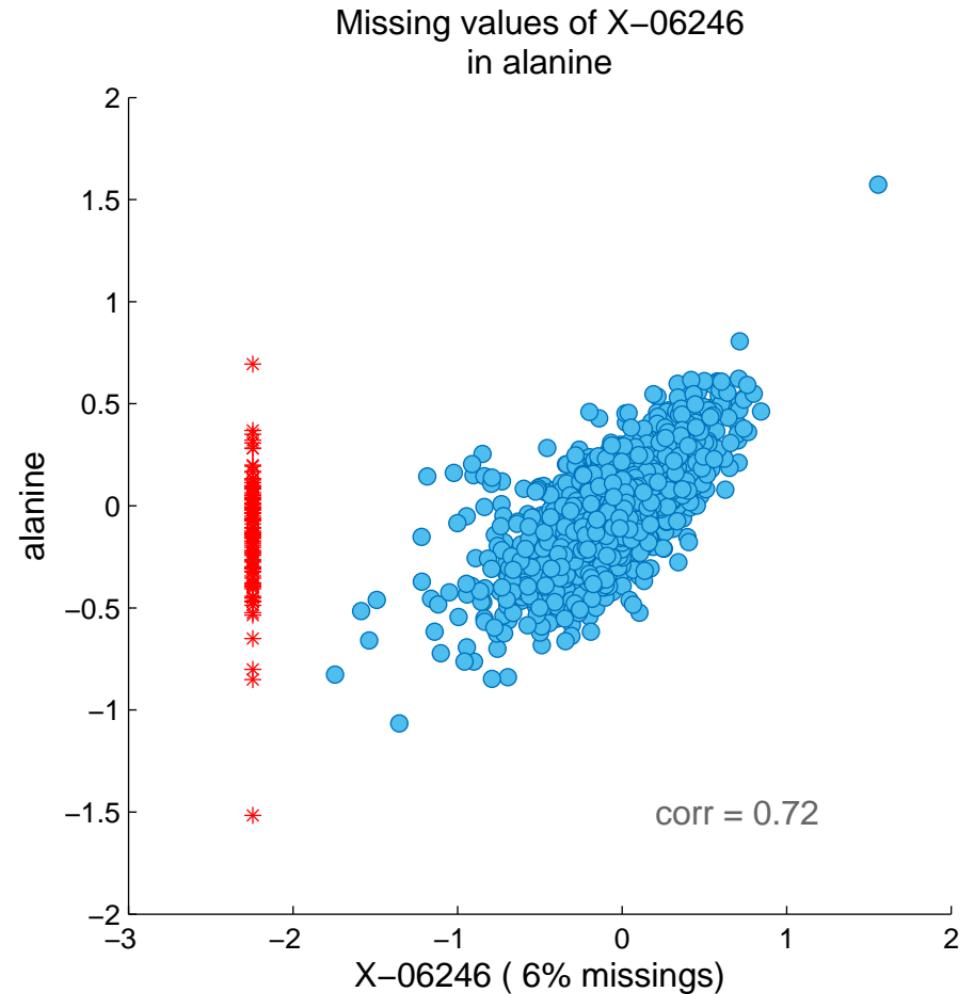


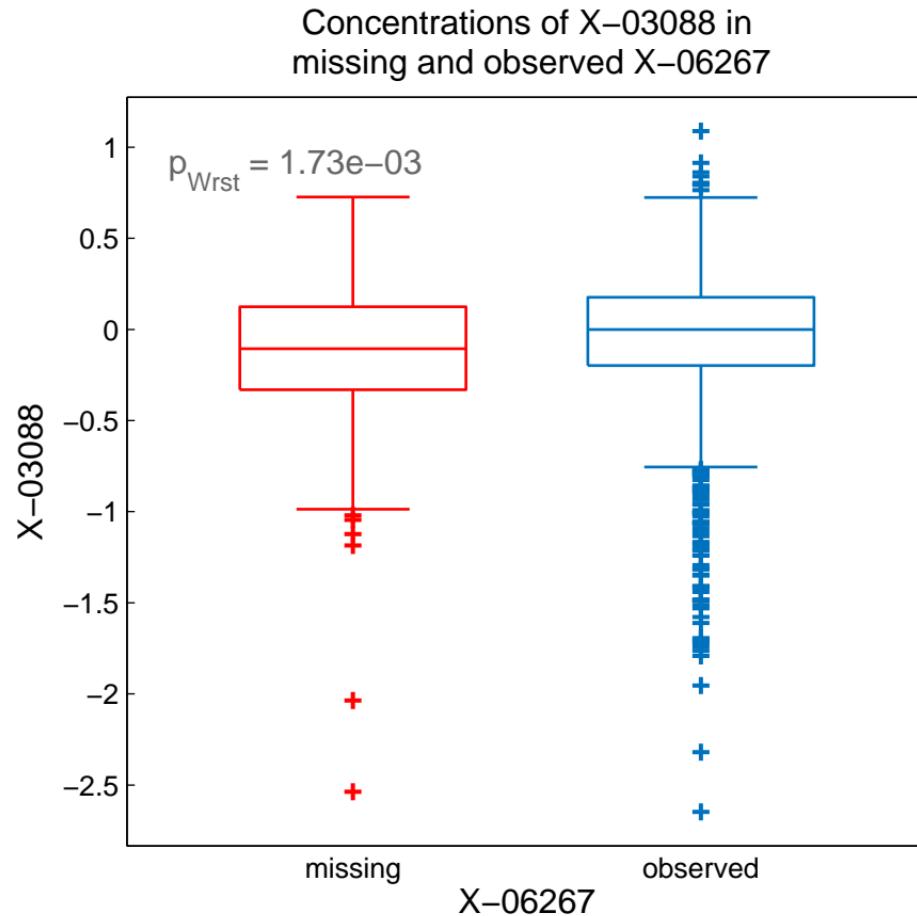
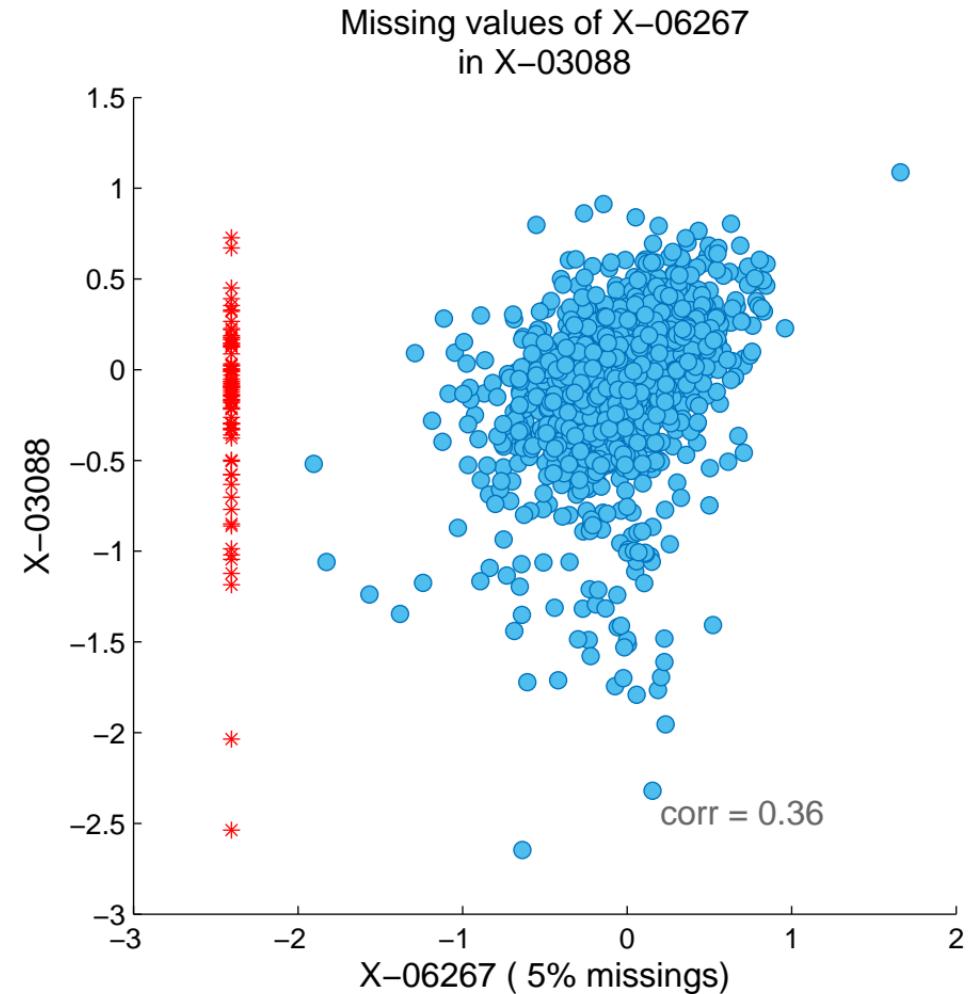


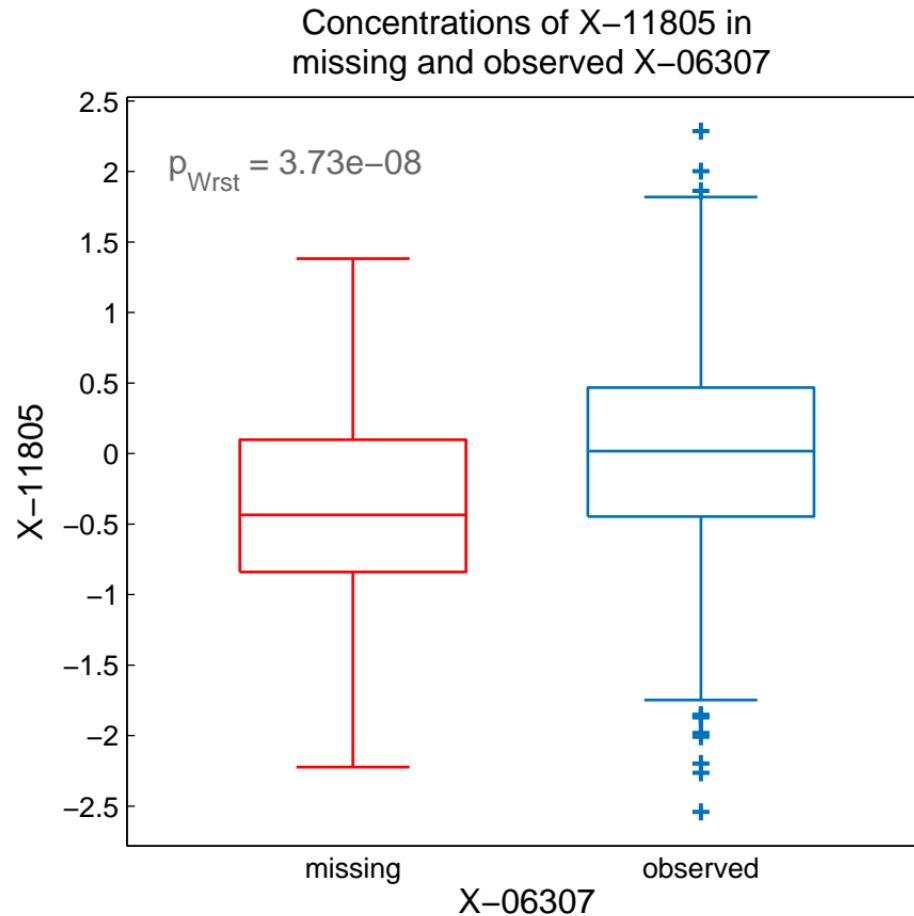
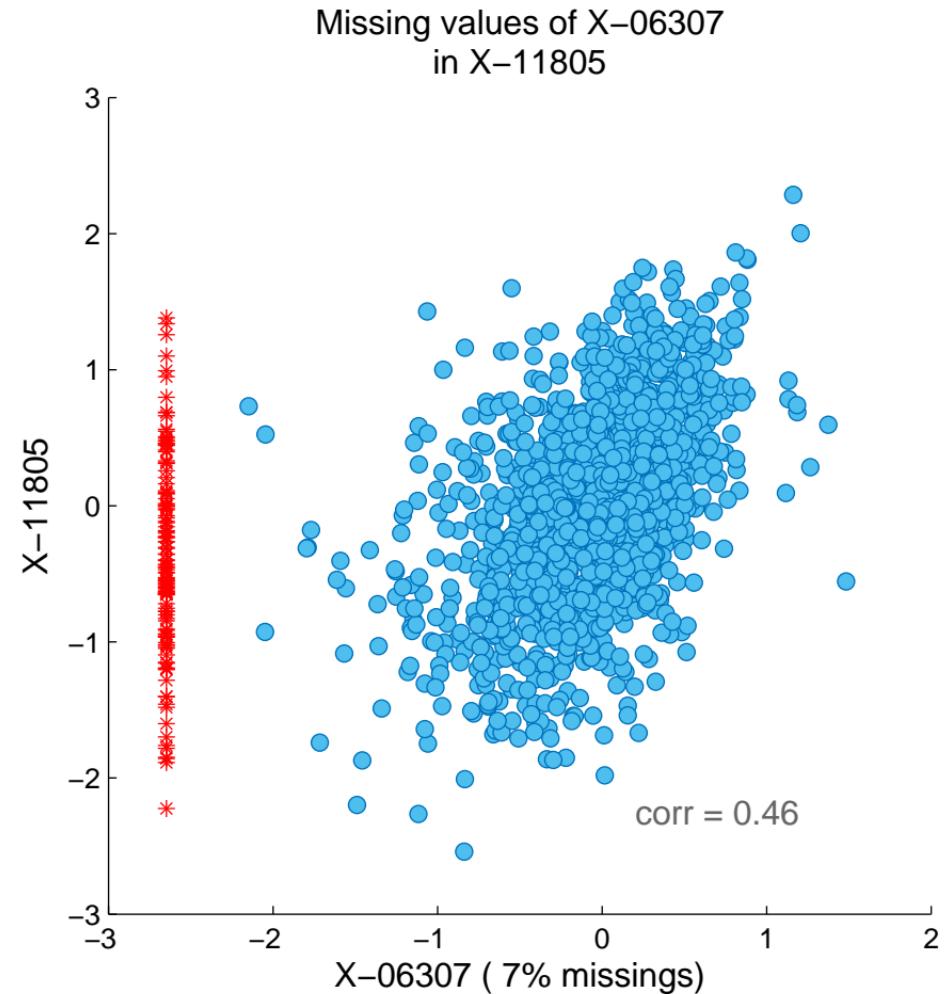


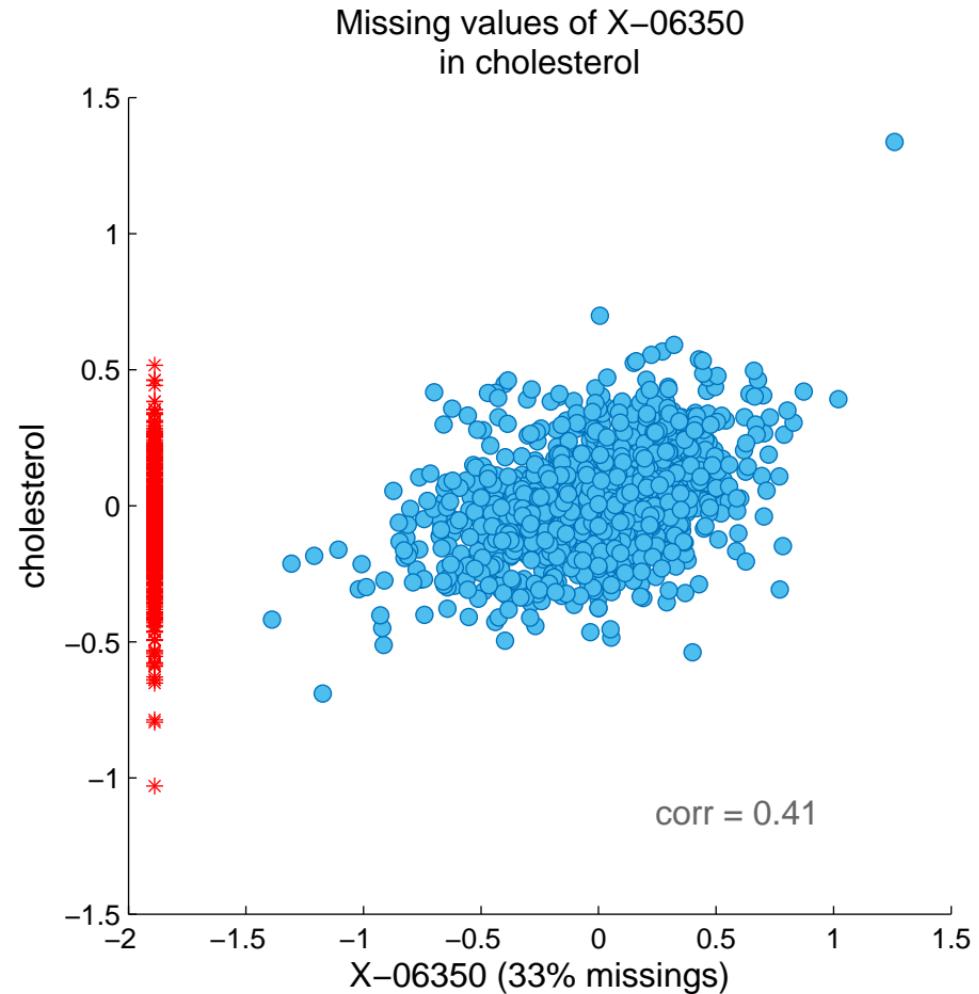




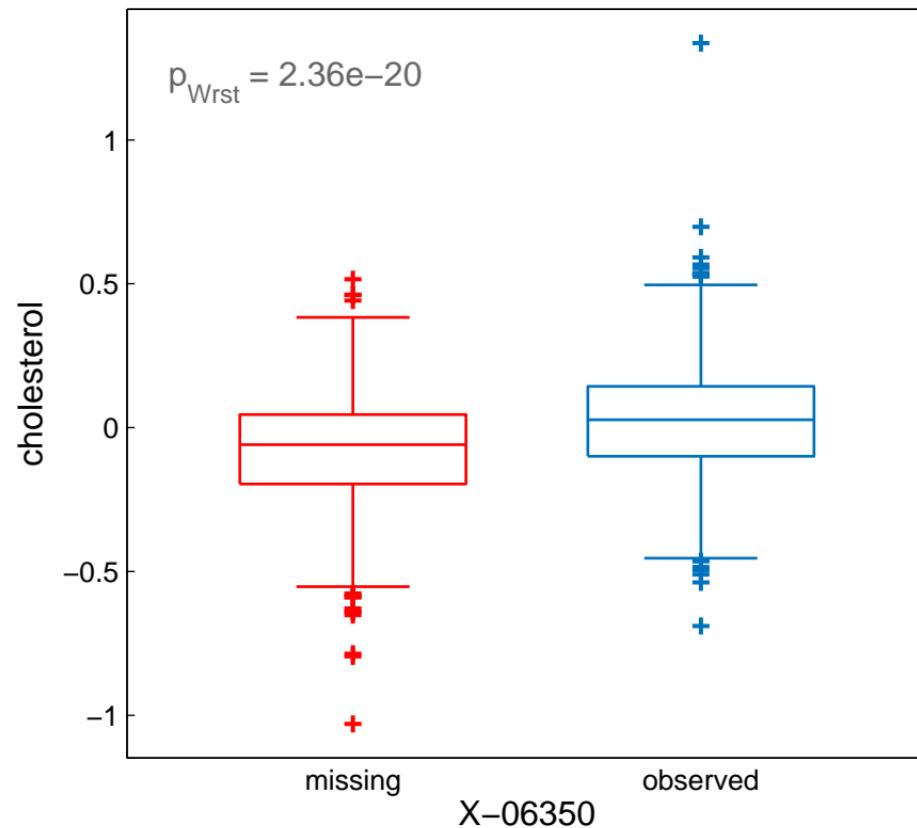


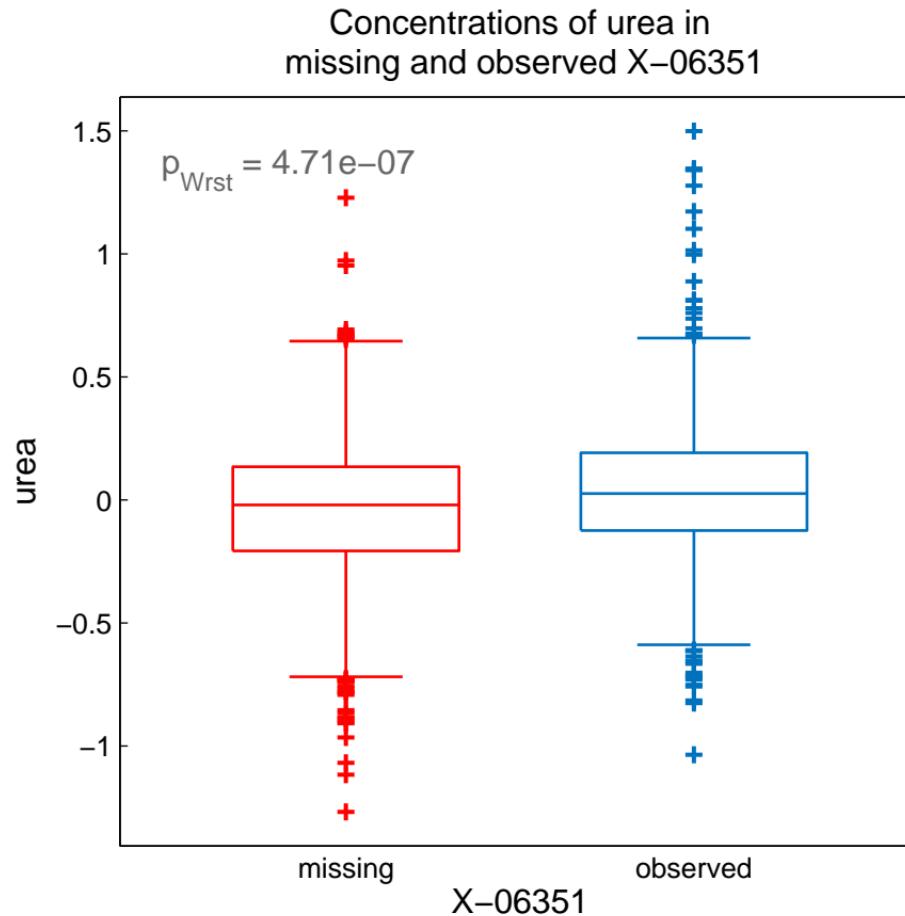
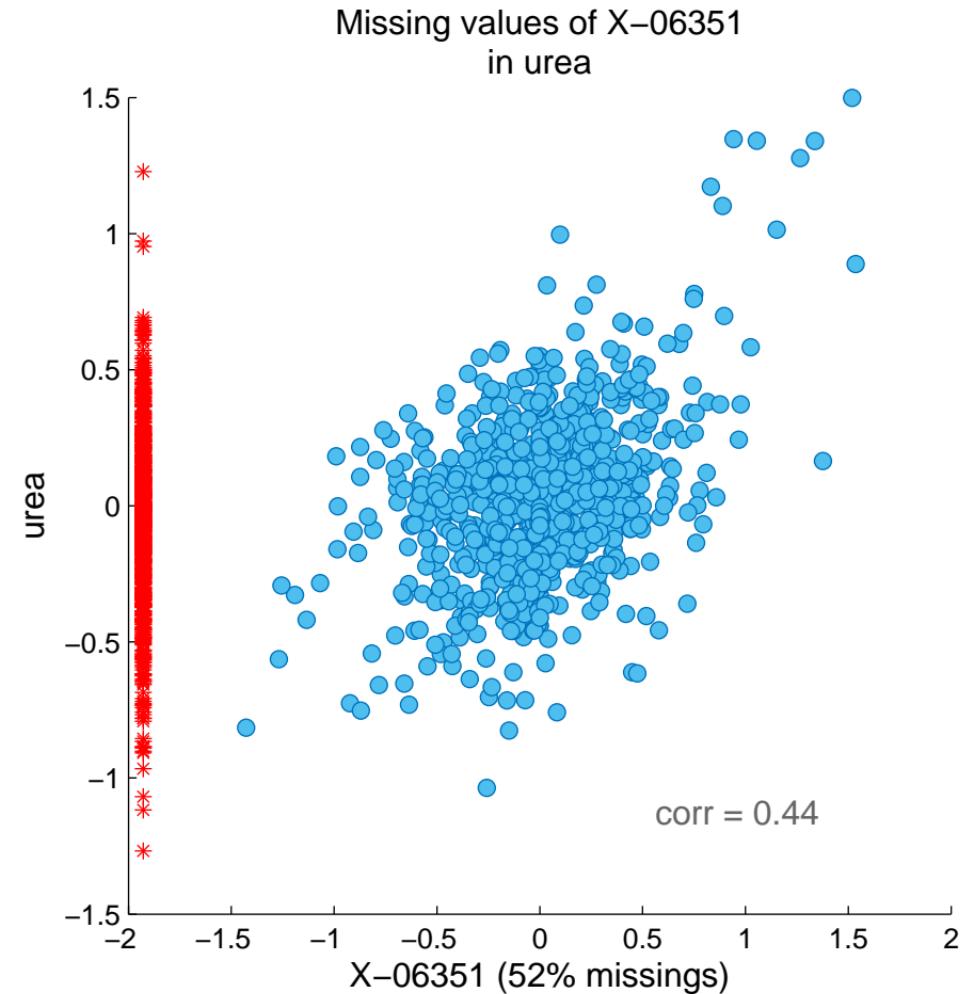


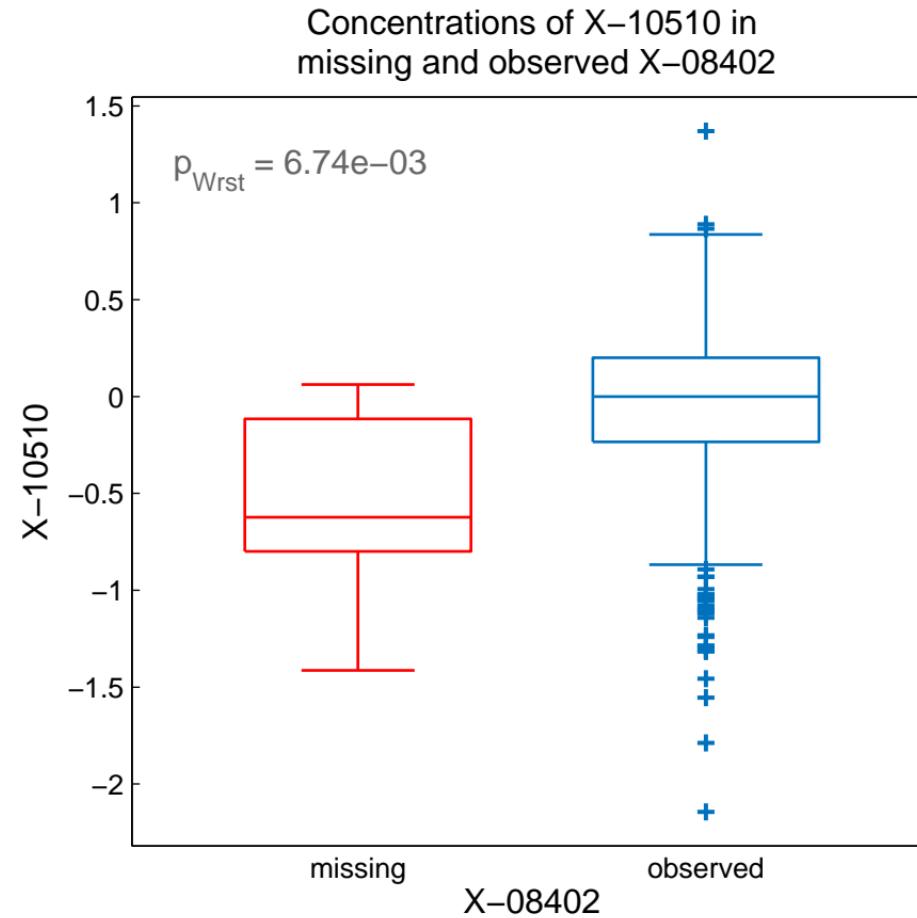
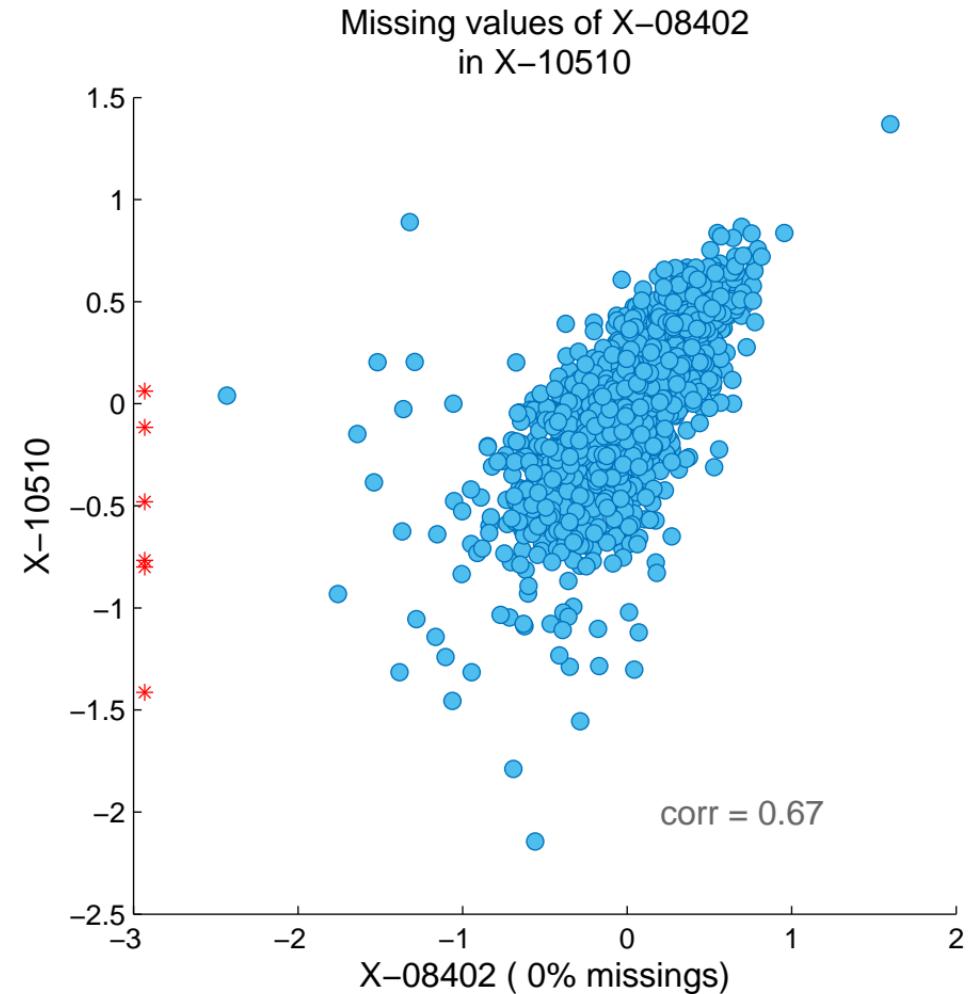


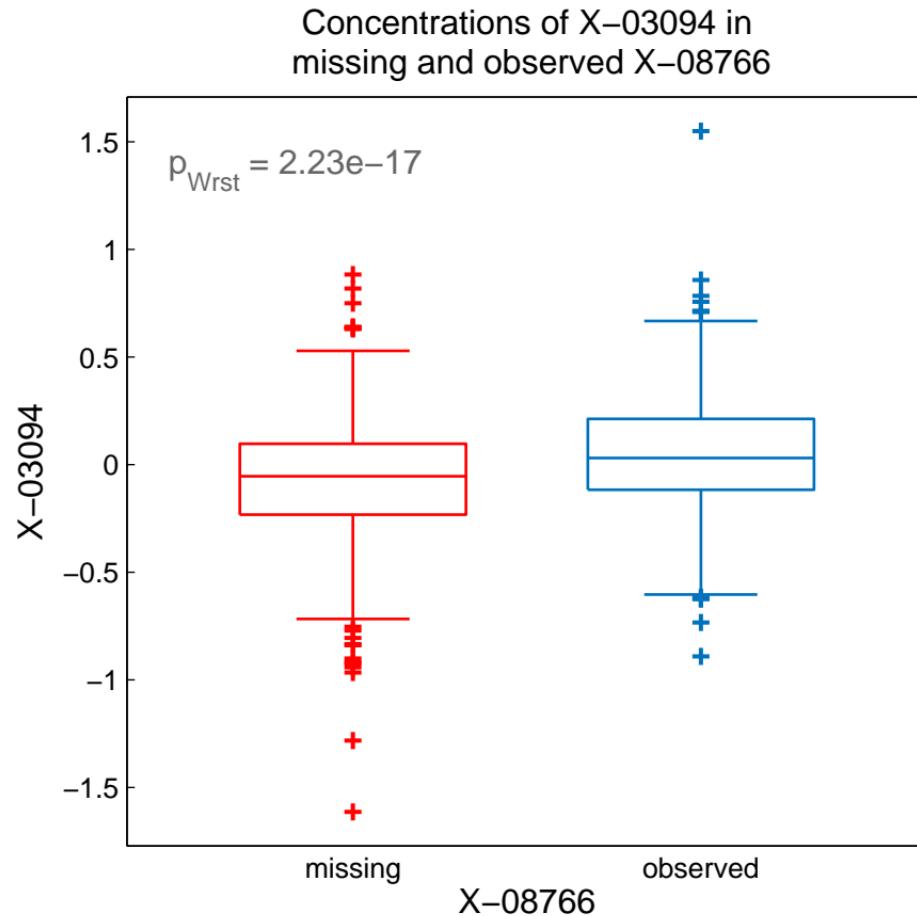
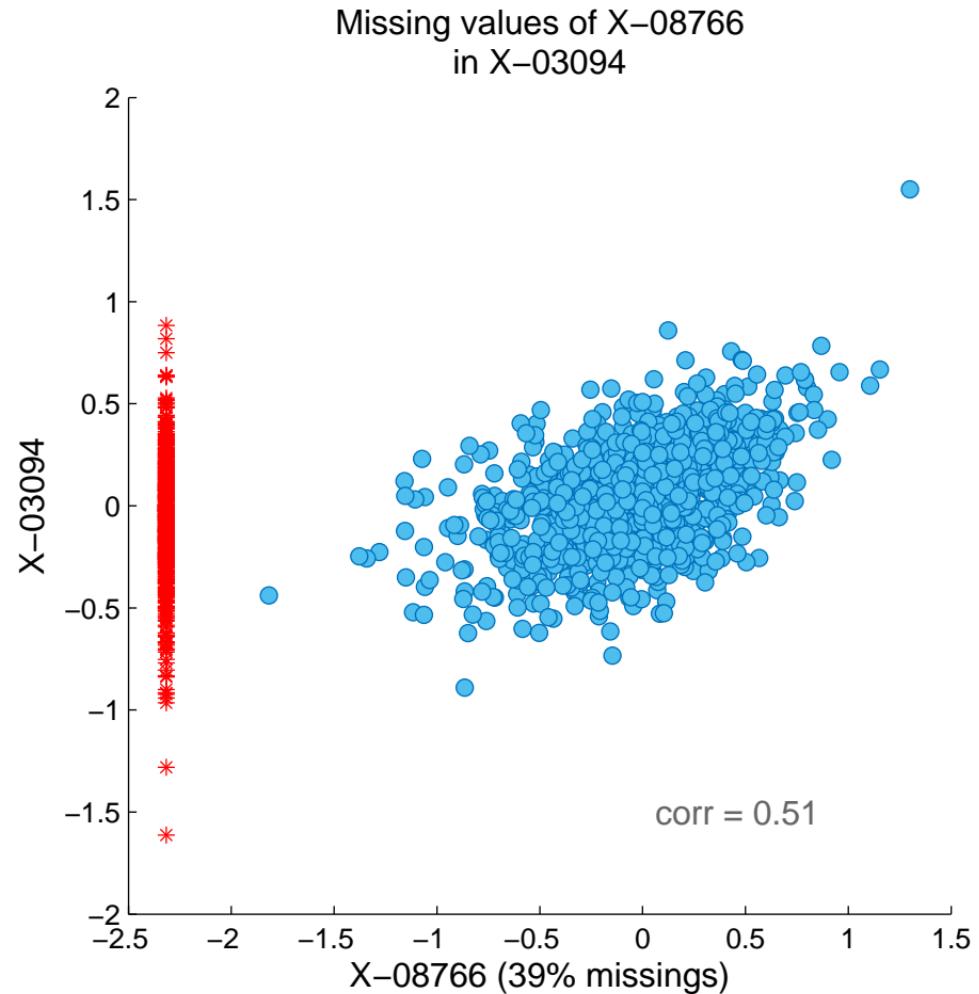


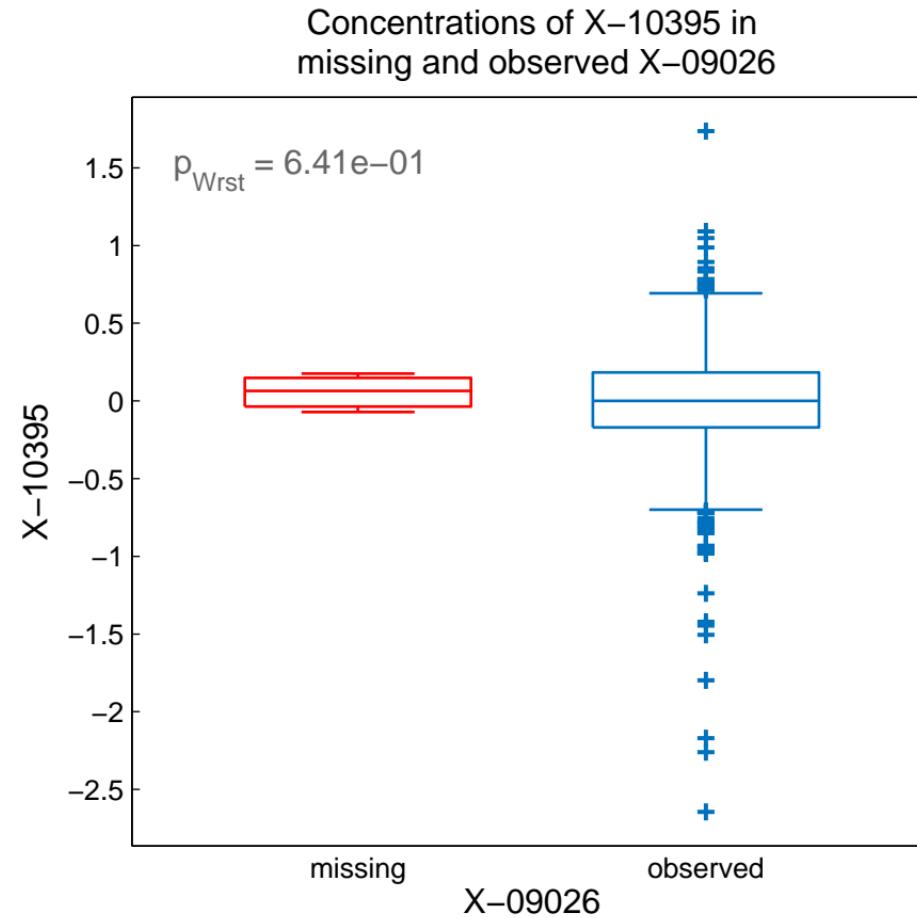
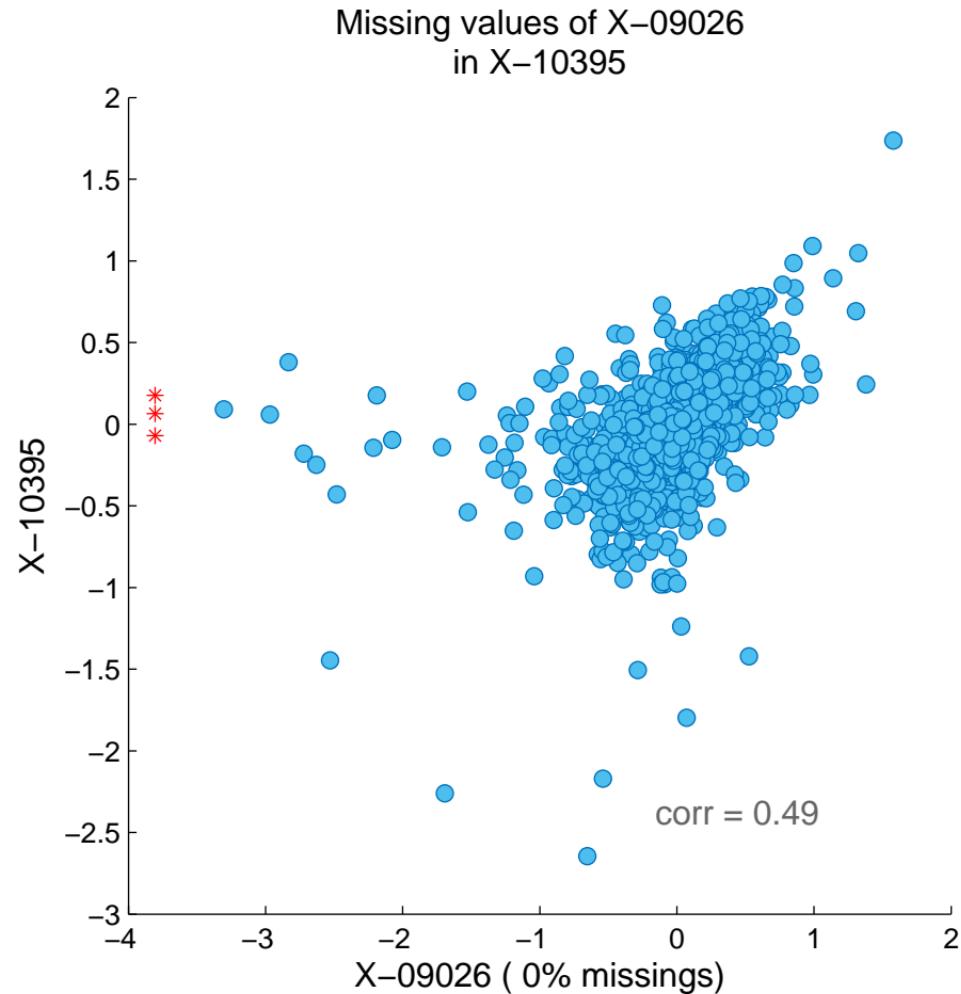
Concentrations of cholesterol in
missing and observed X-06350

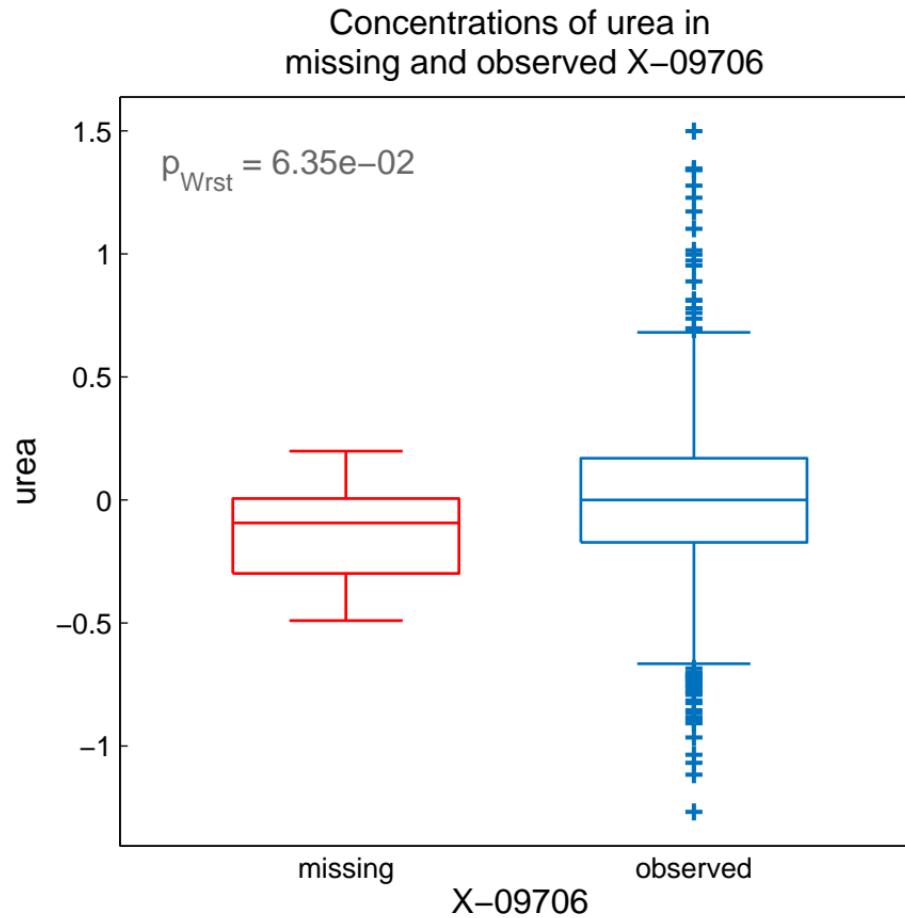
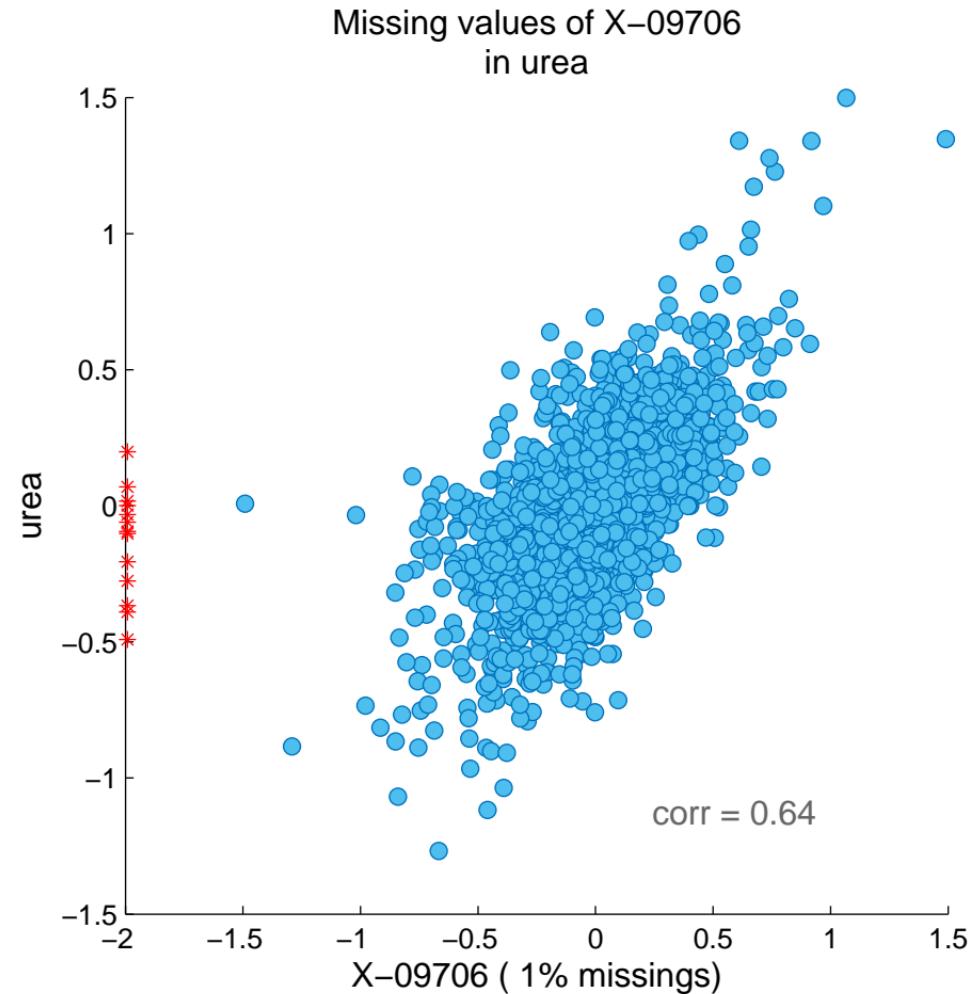


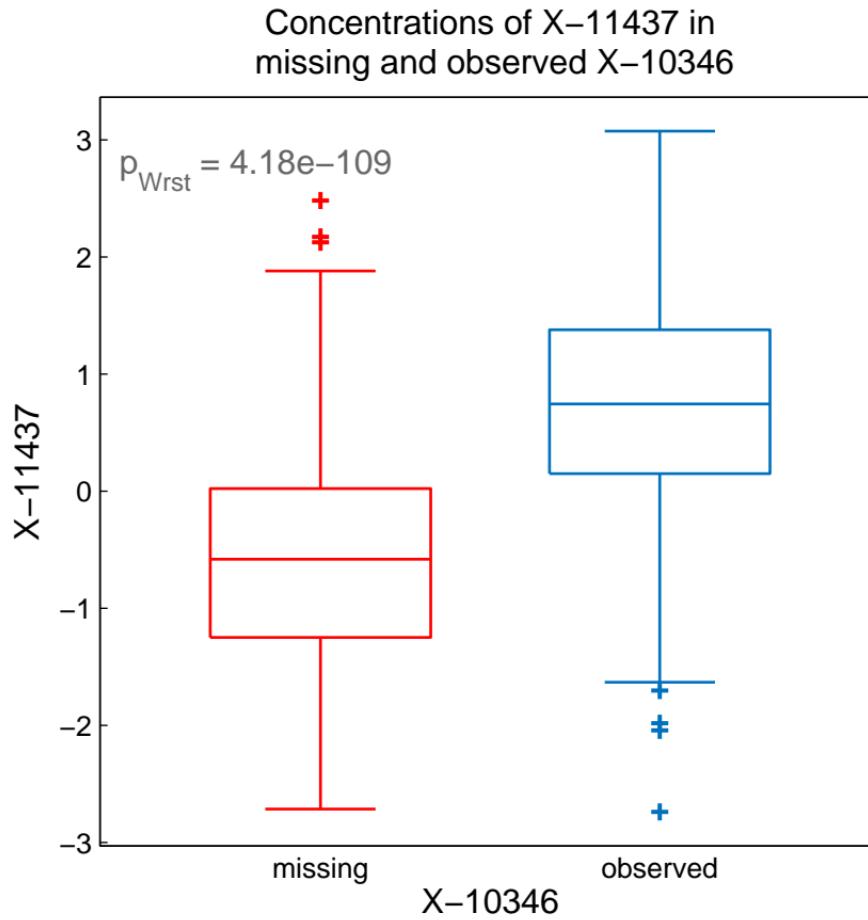
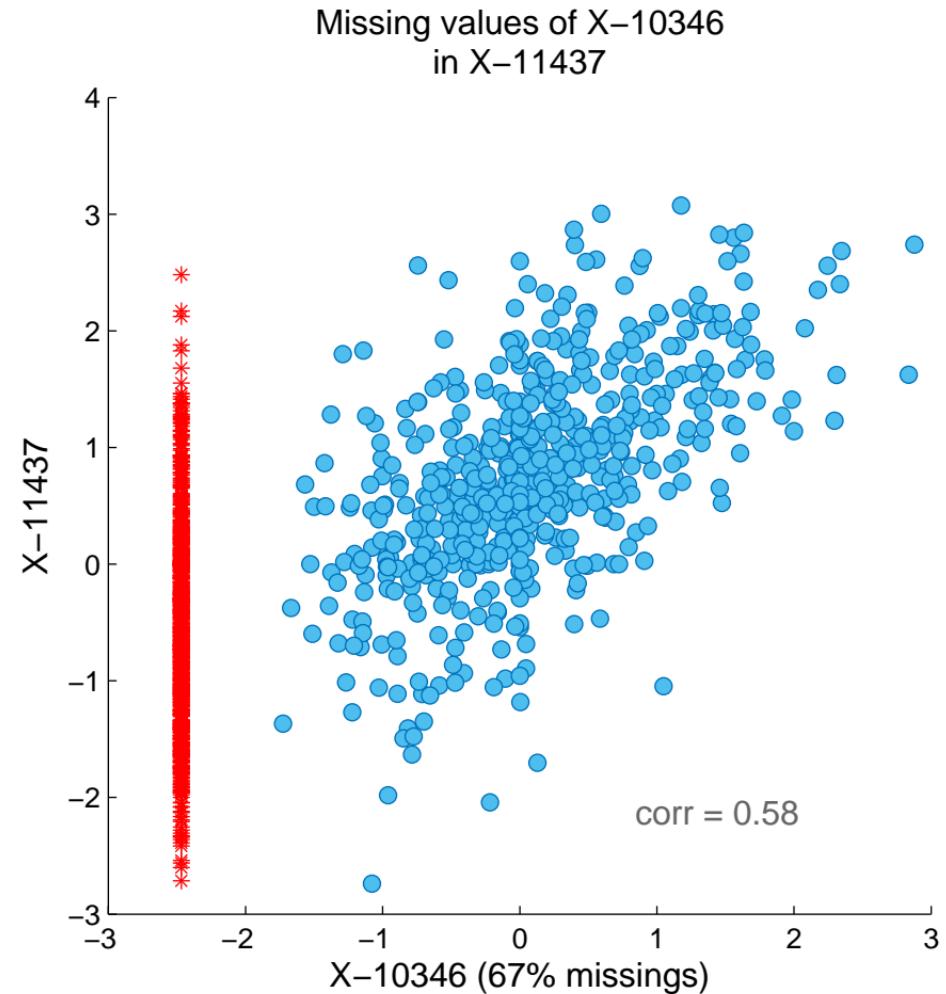


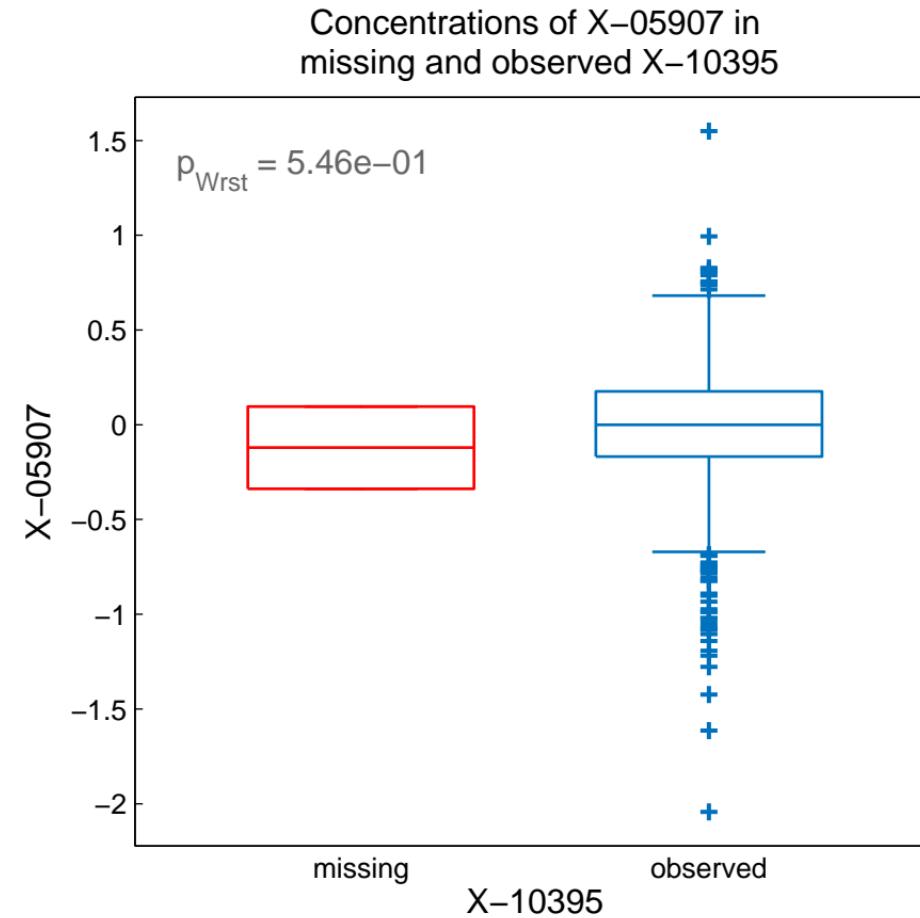
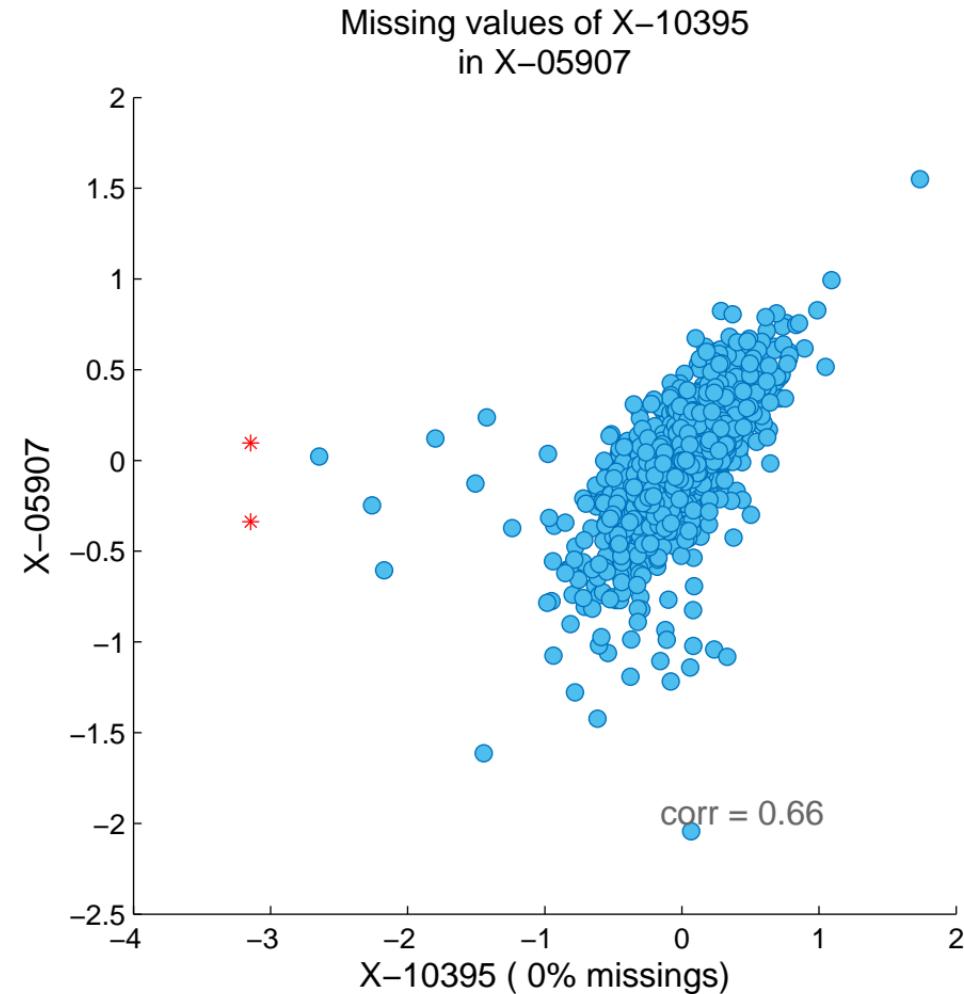


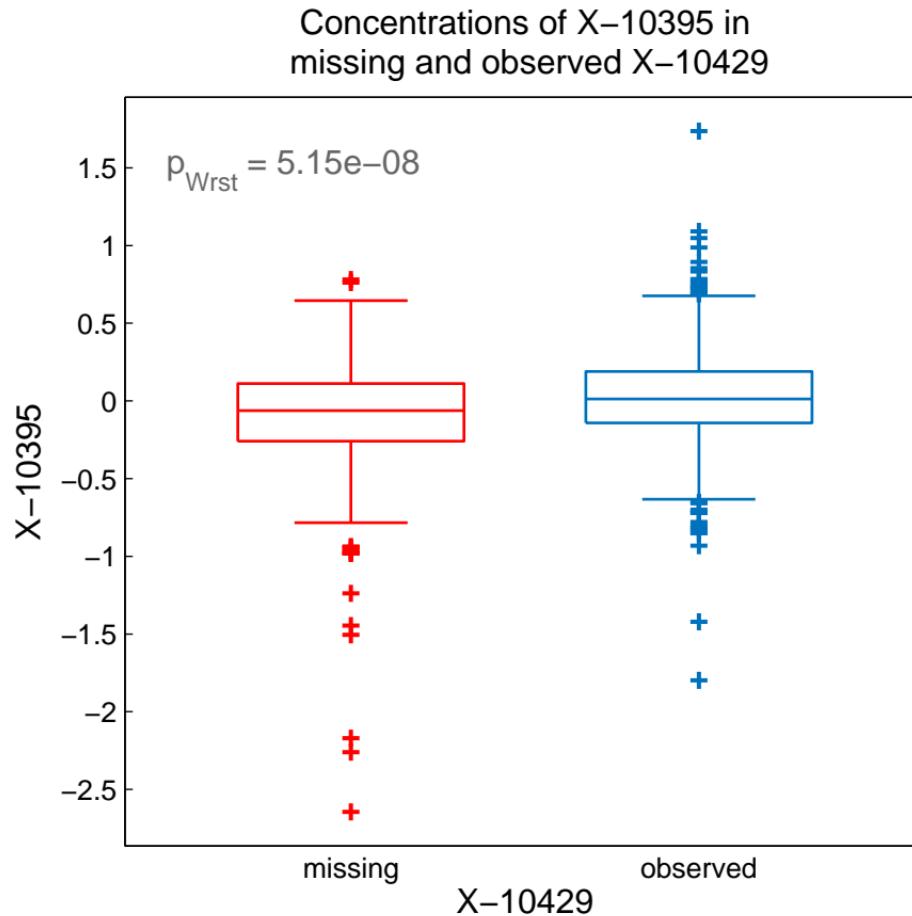
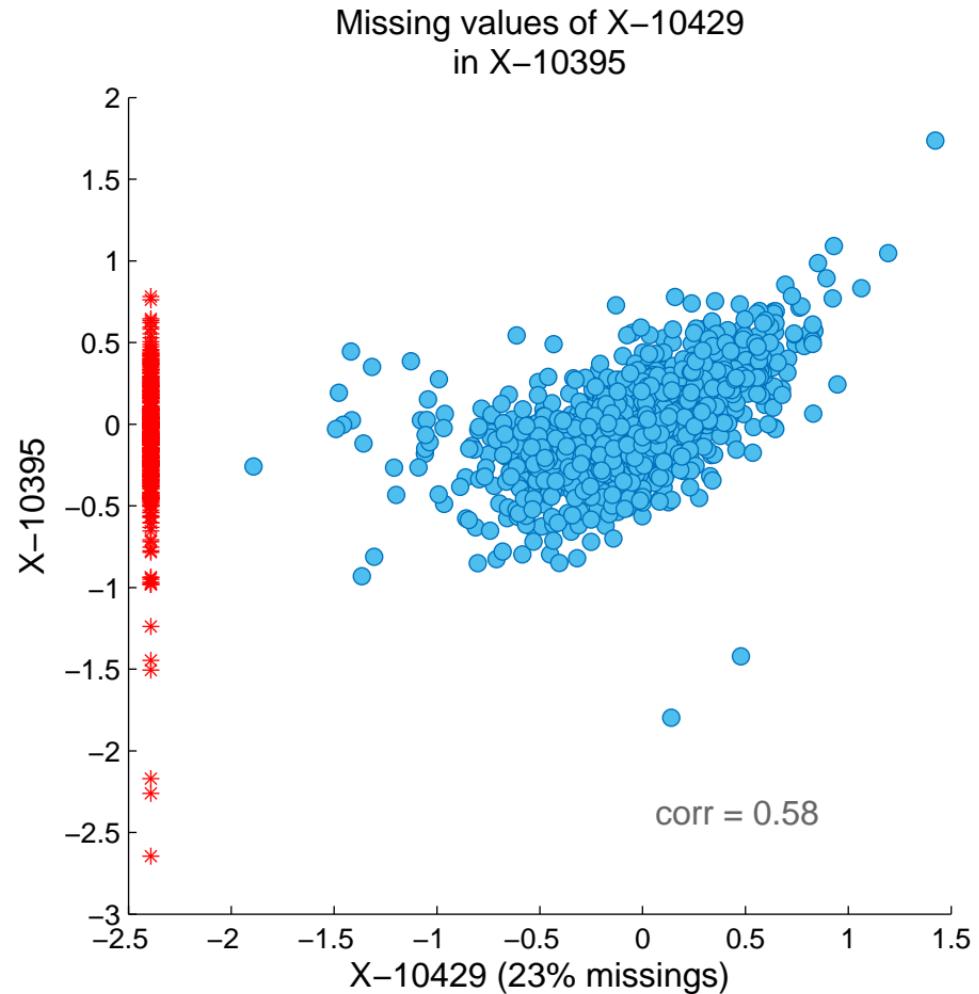


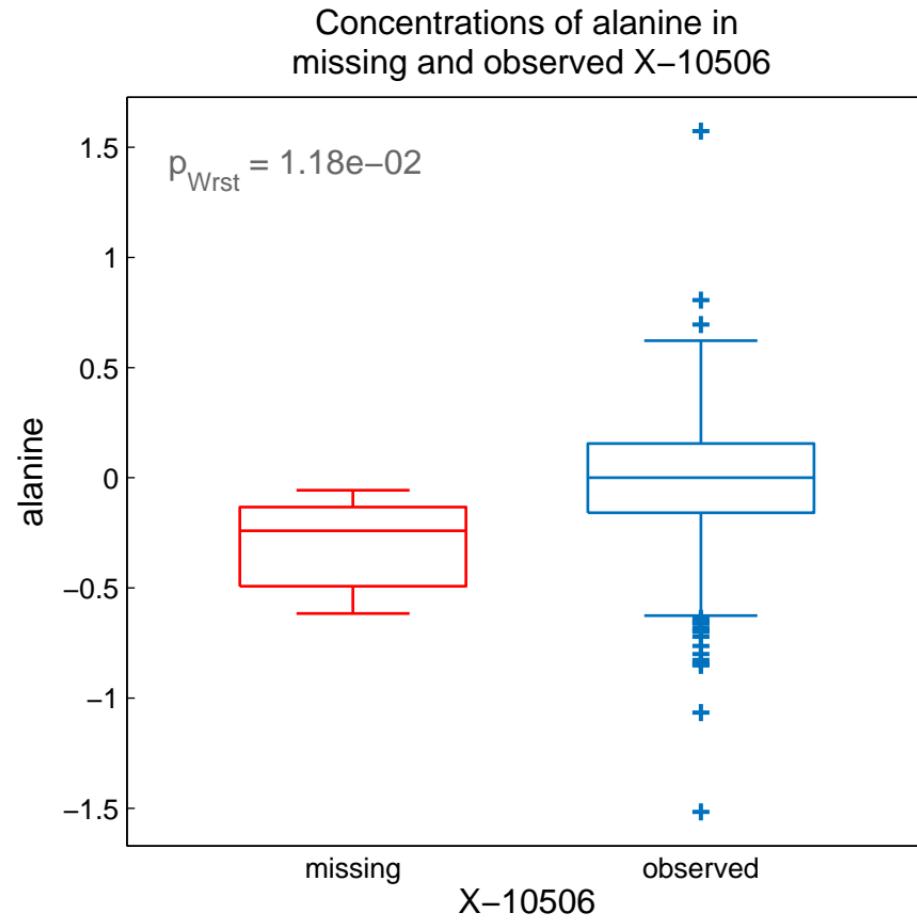
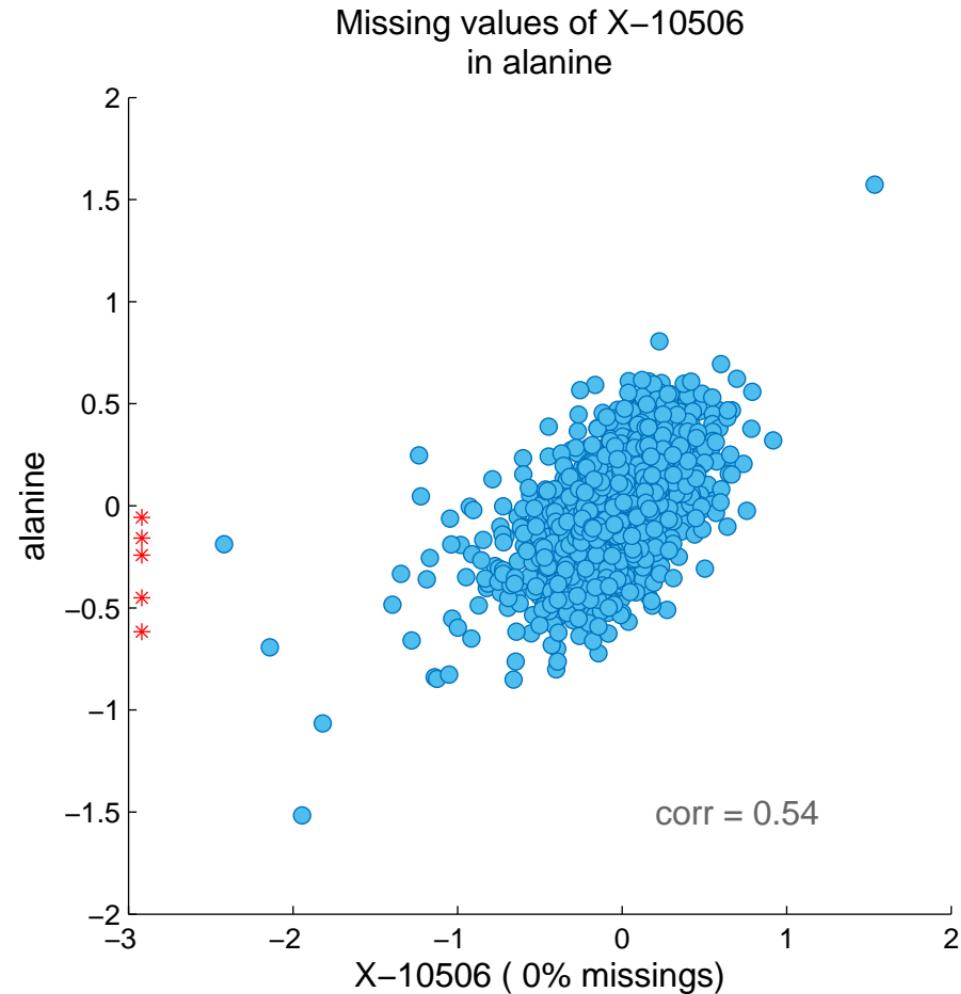


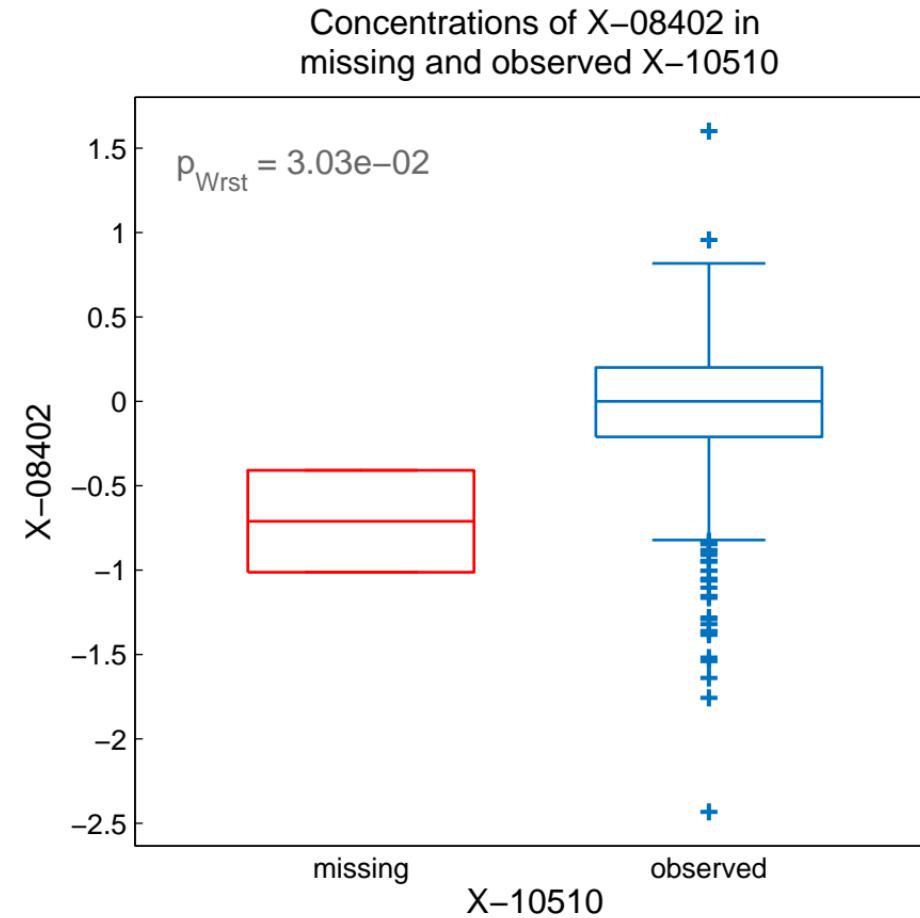
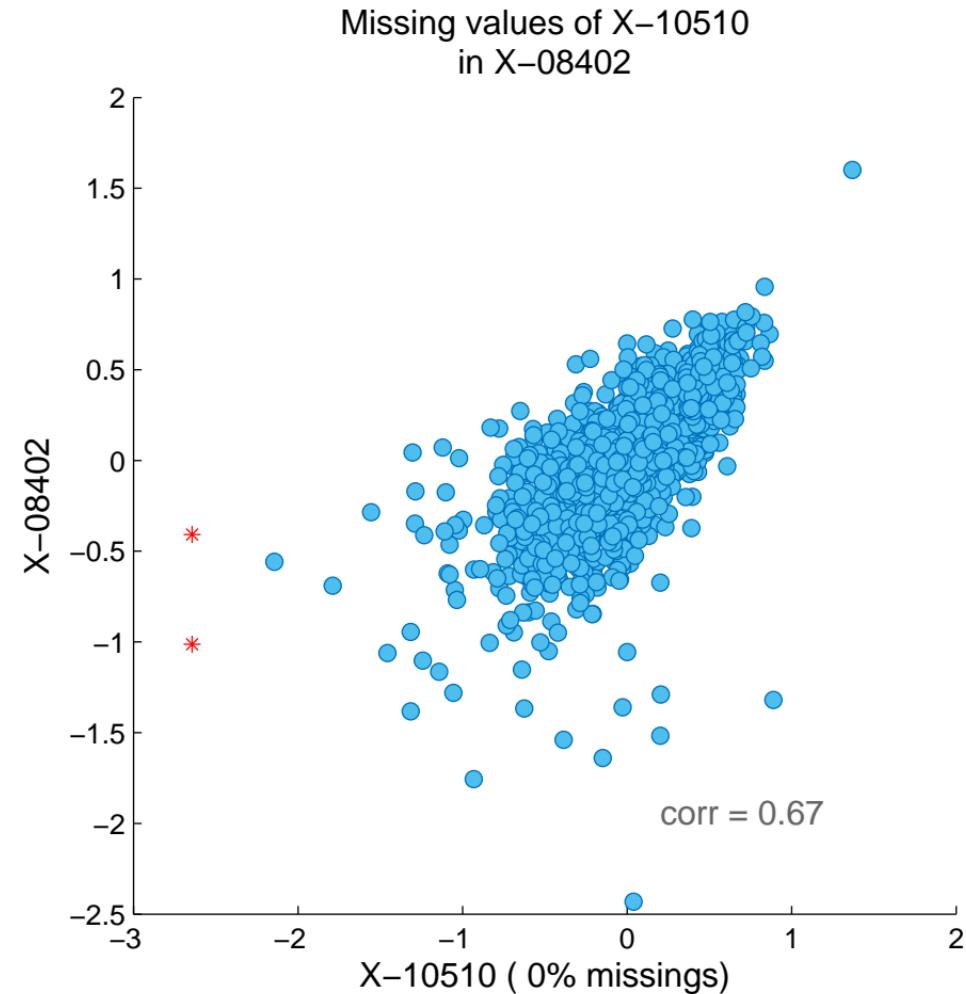


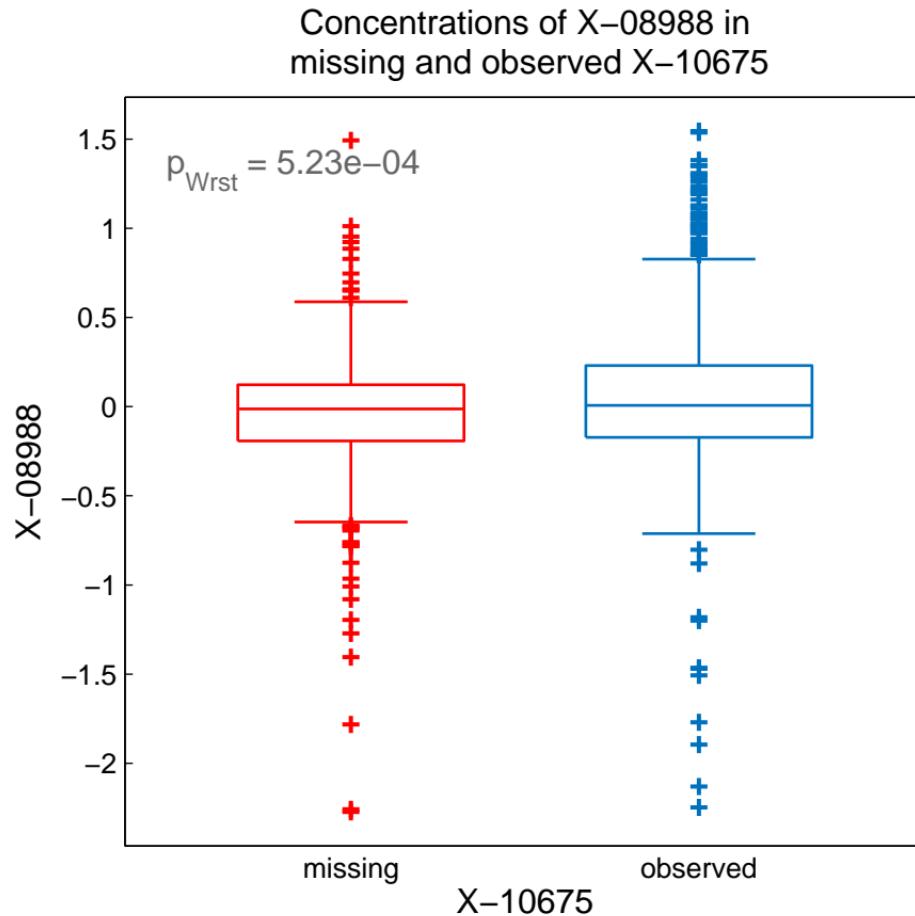
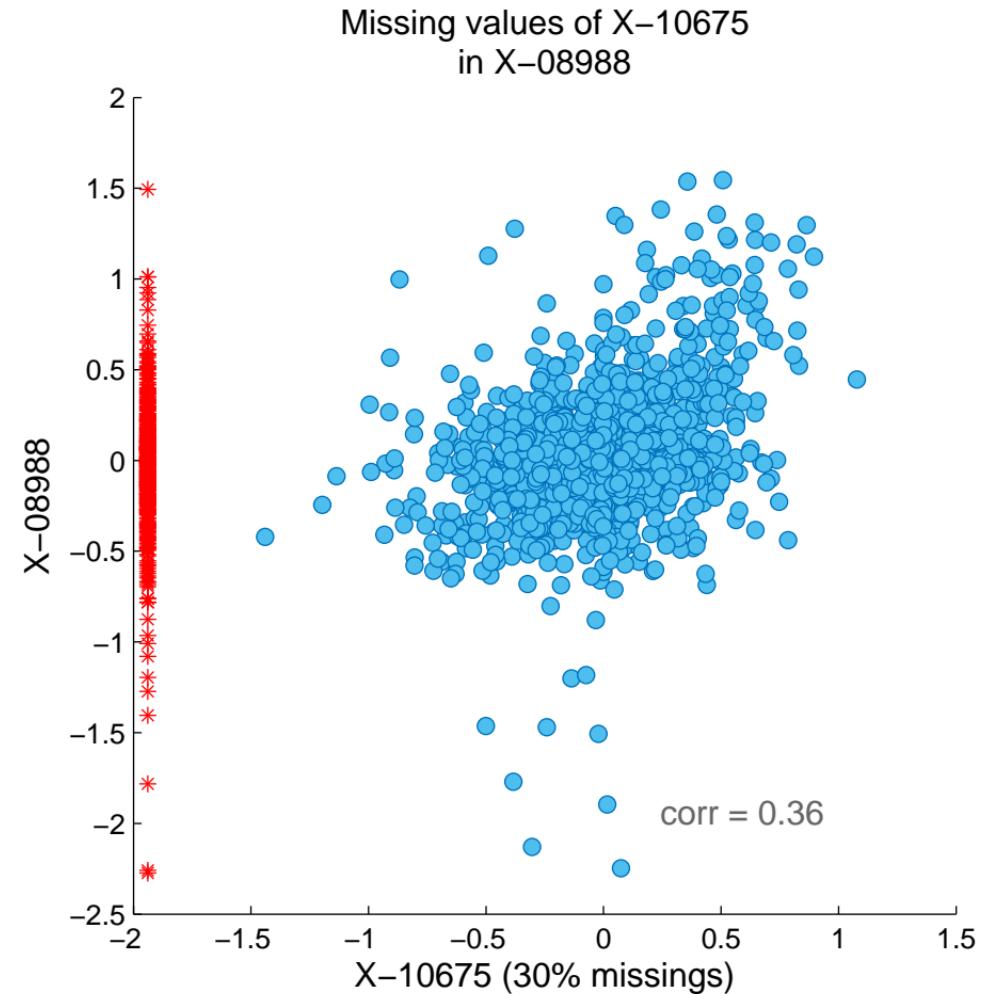


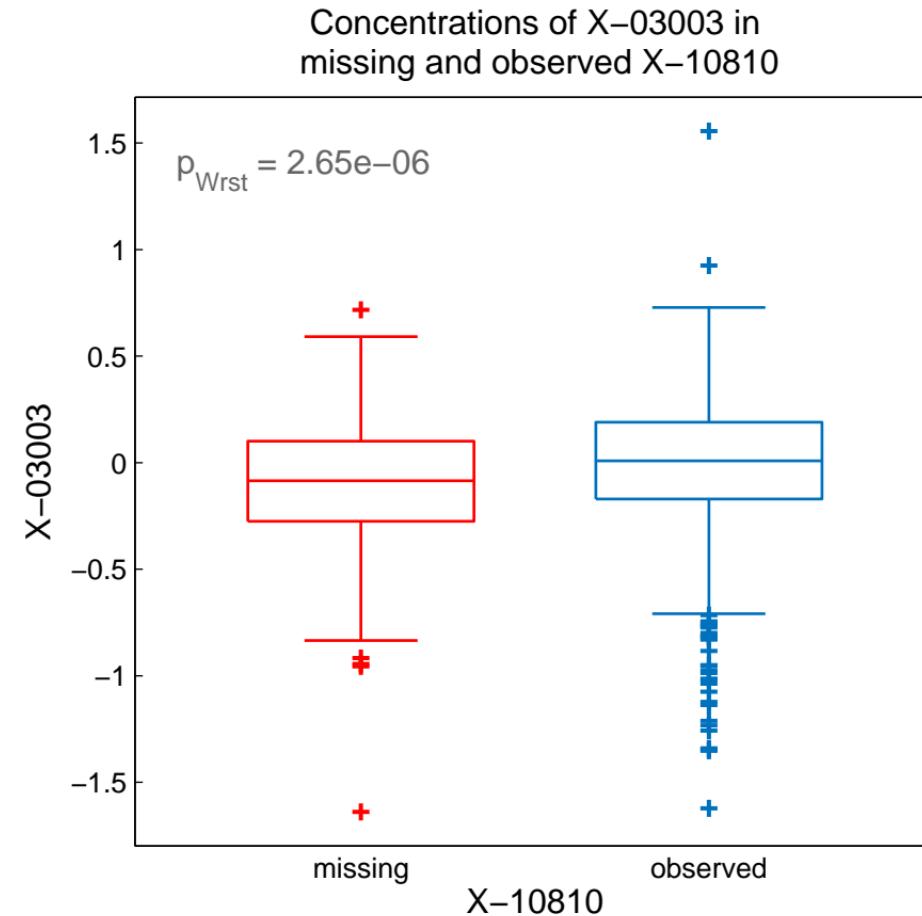
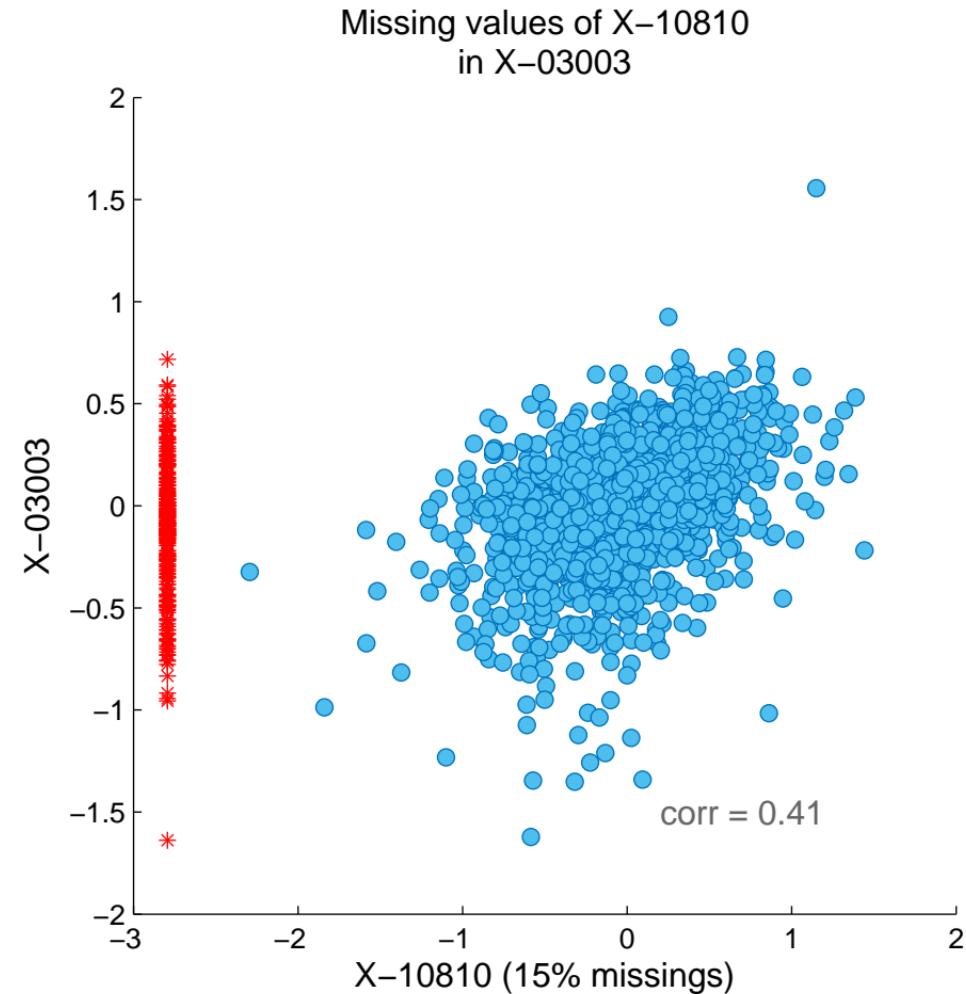


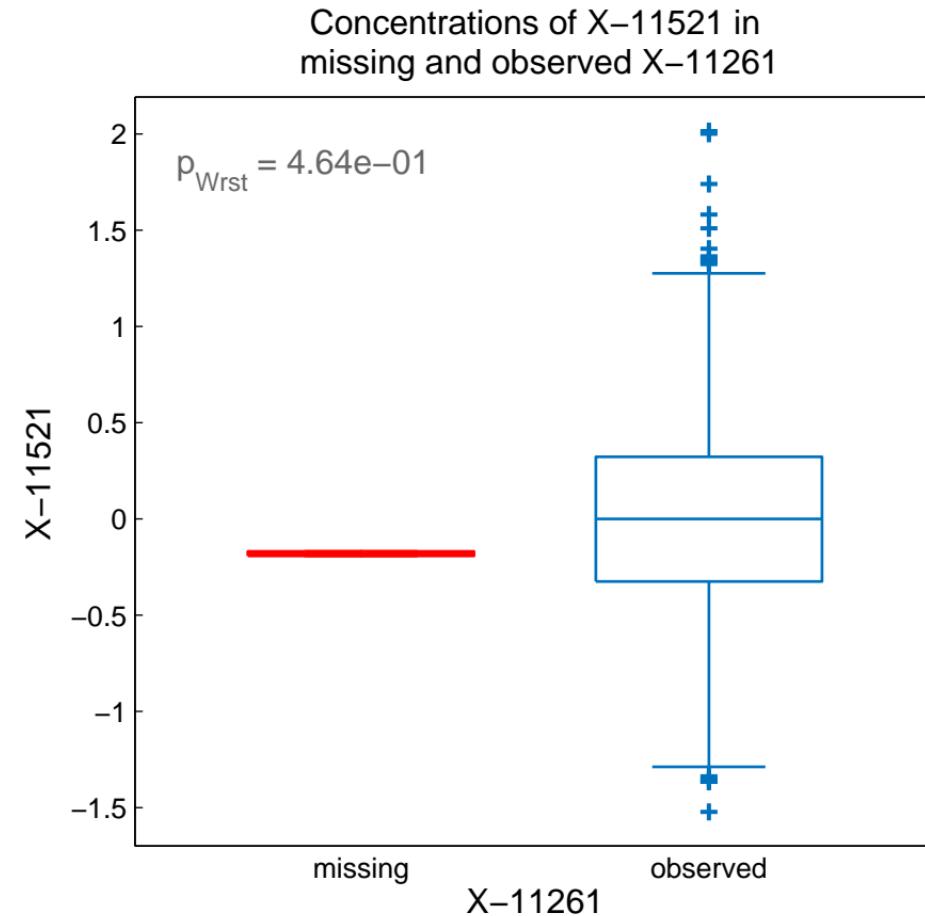
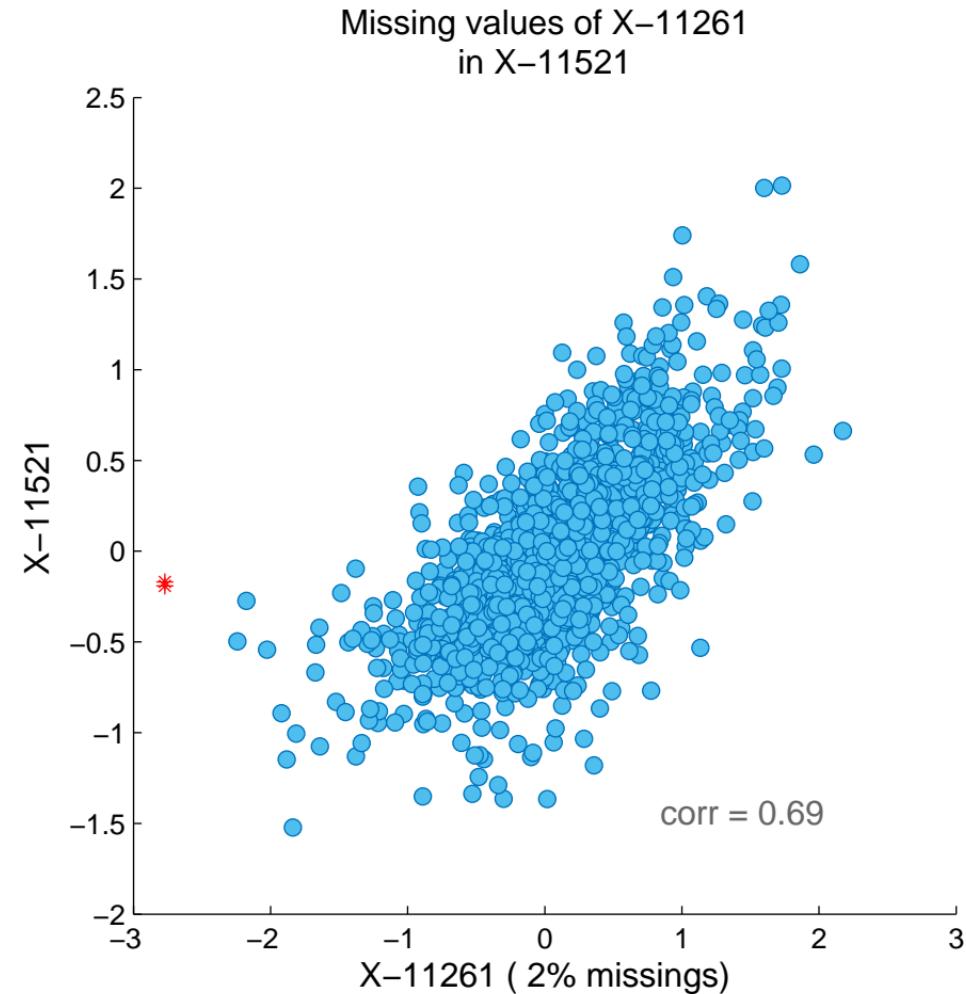


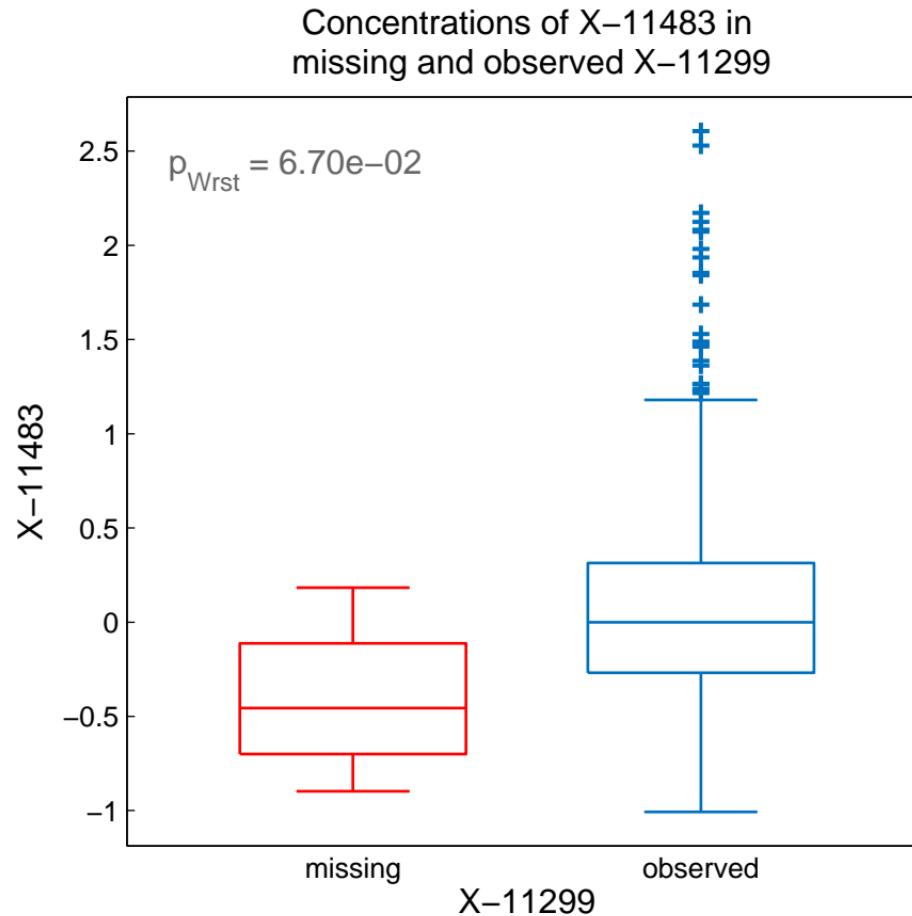
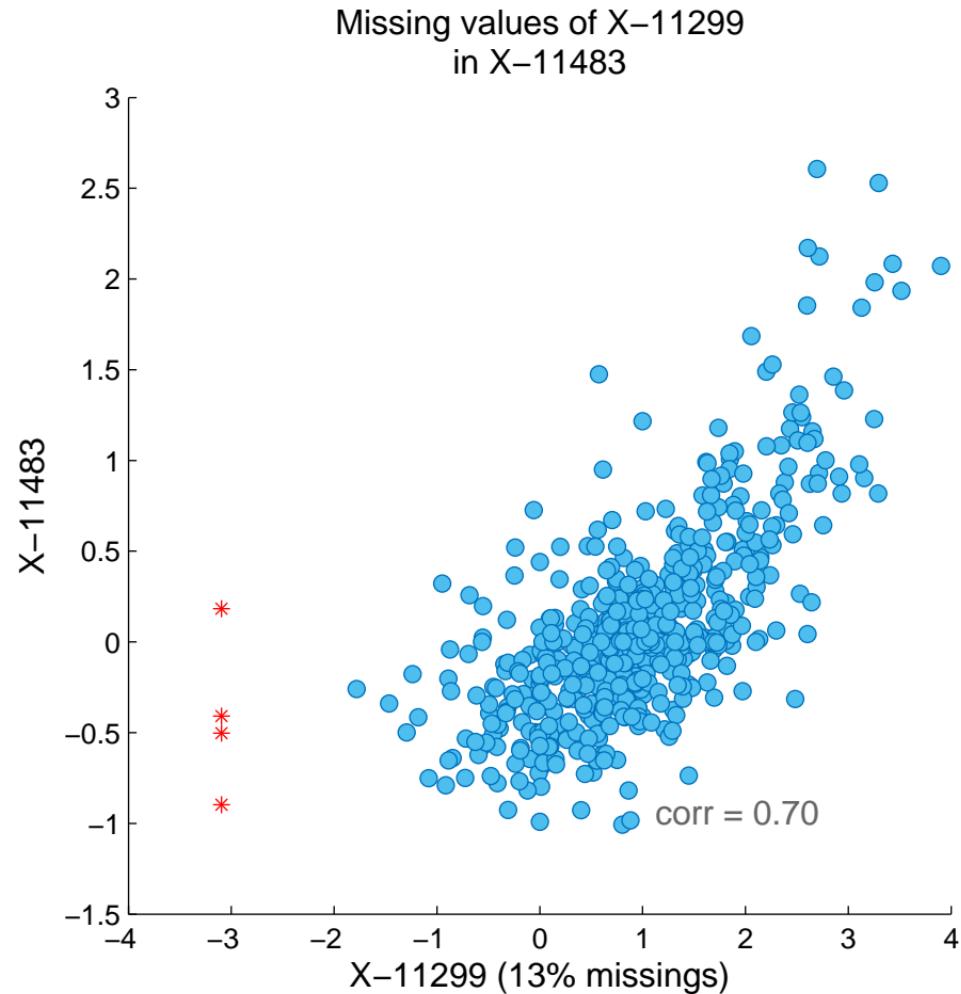




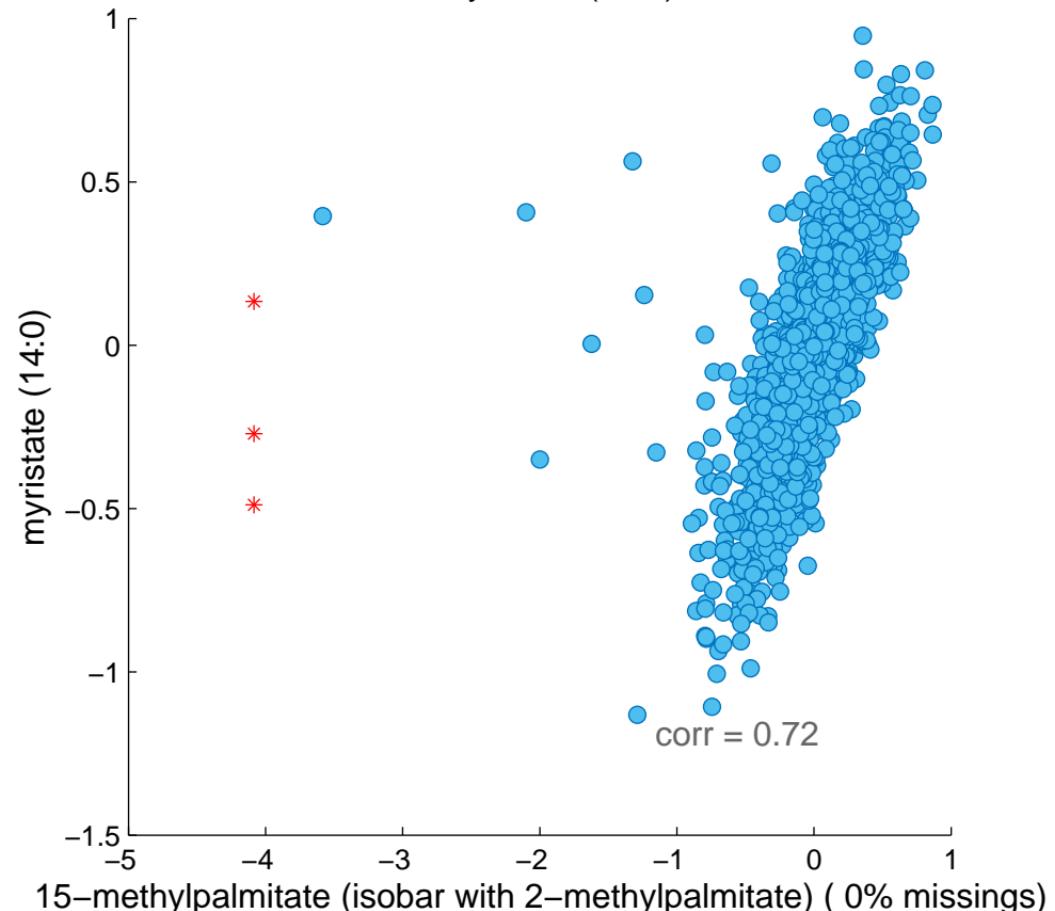




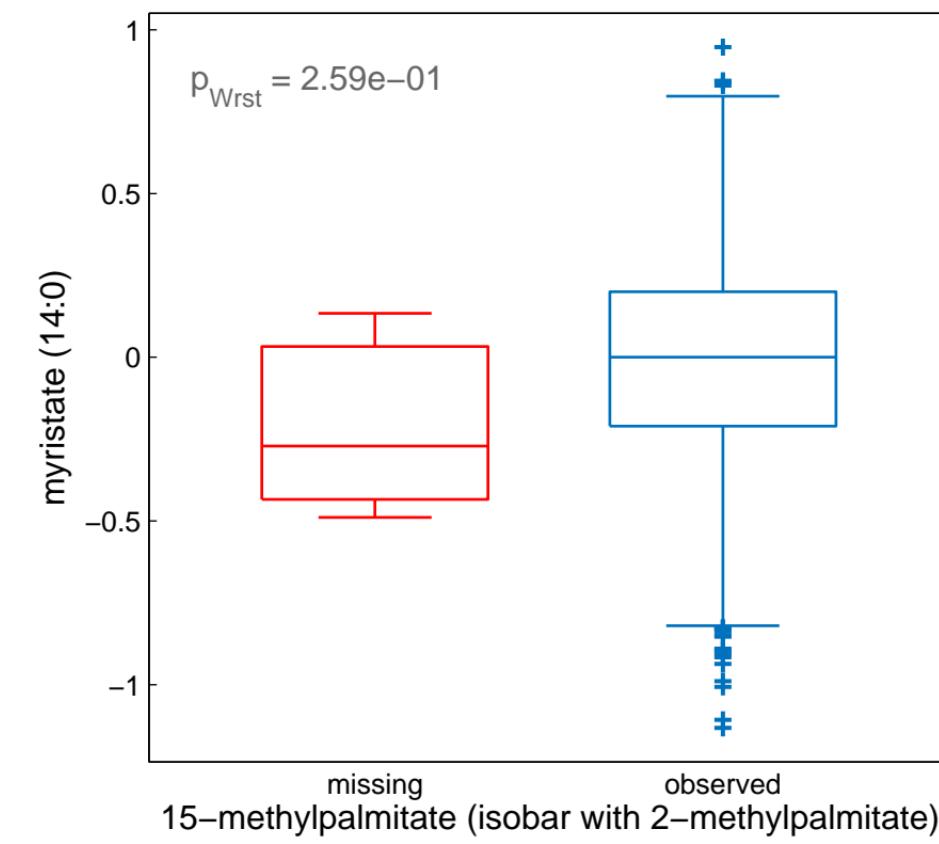


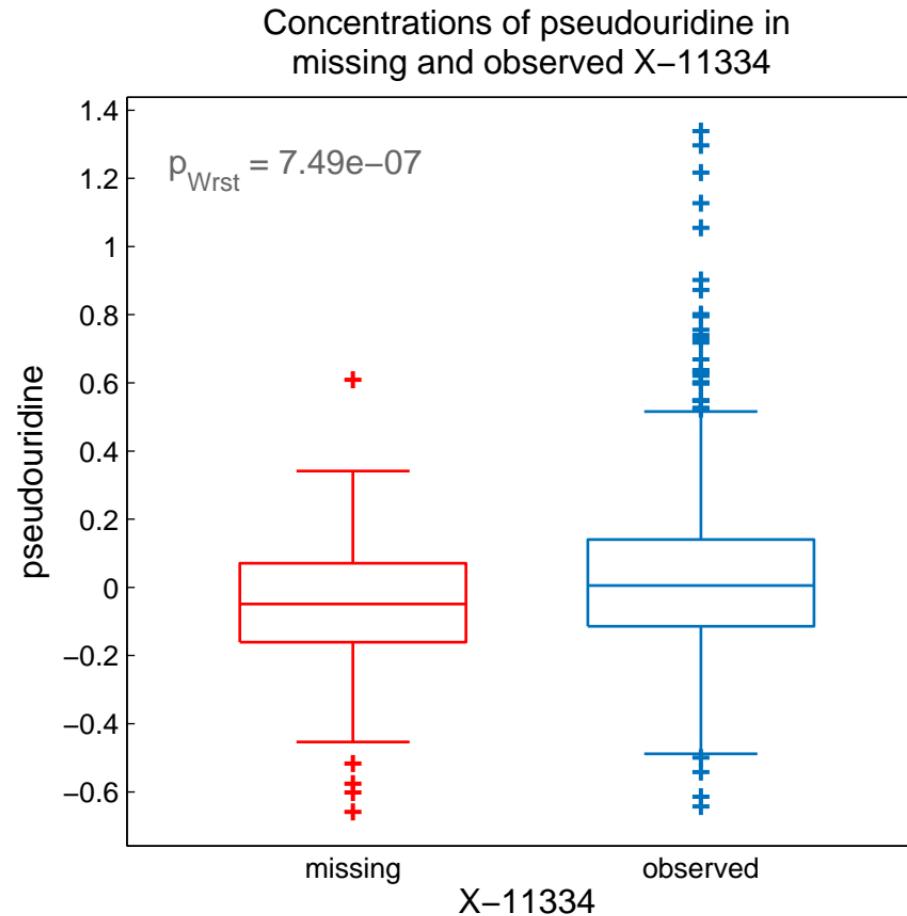
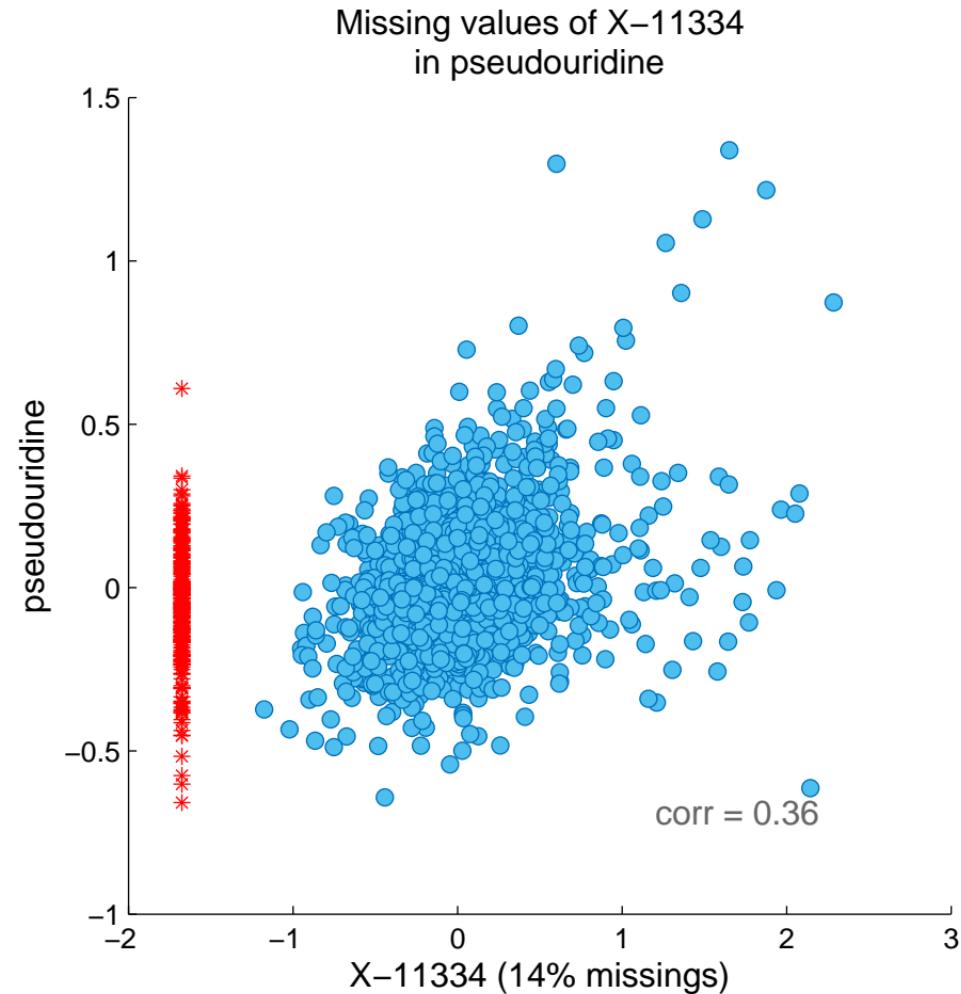


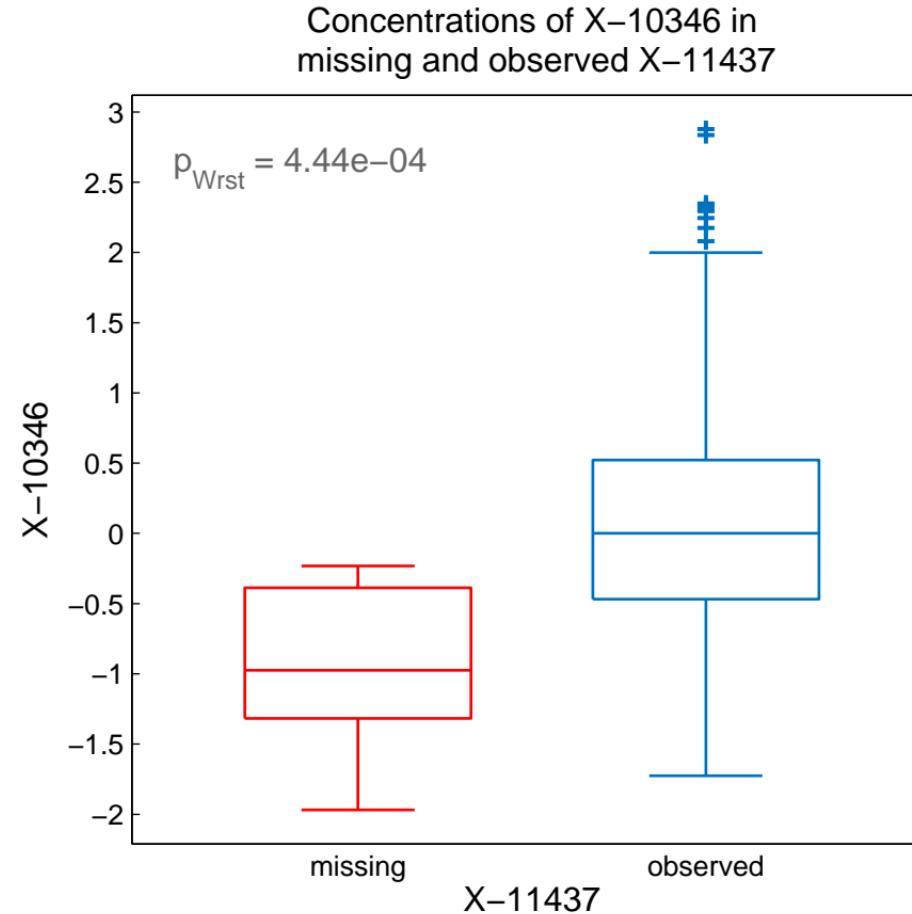
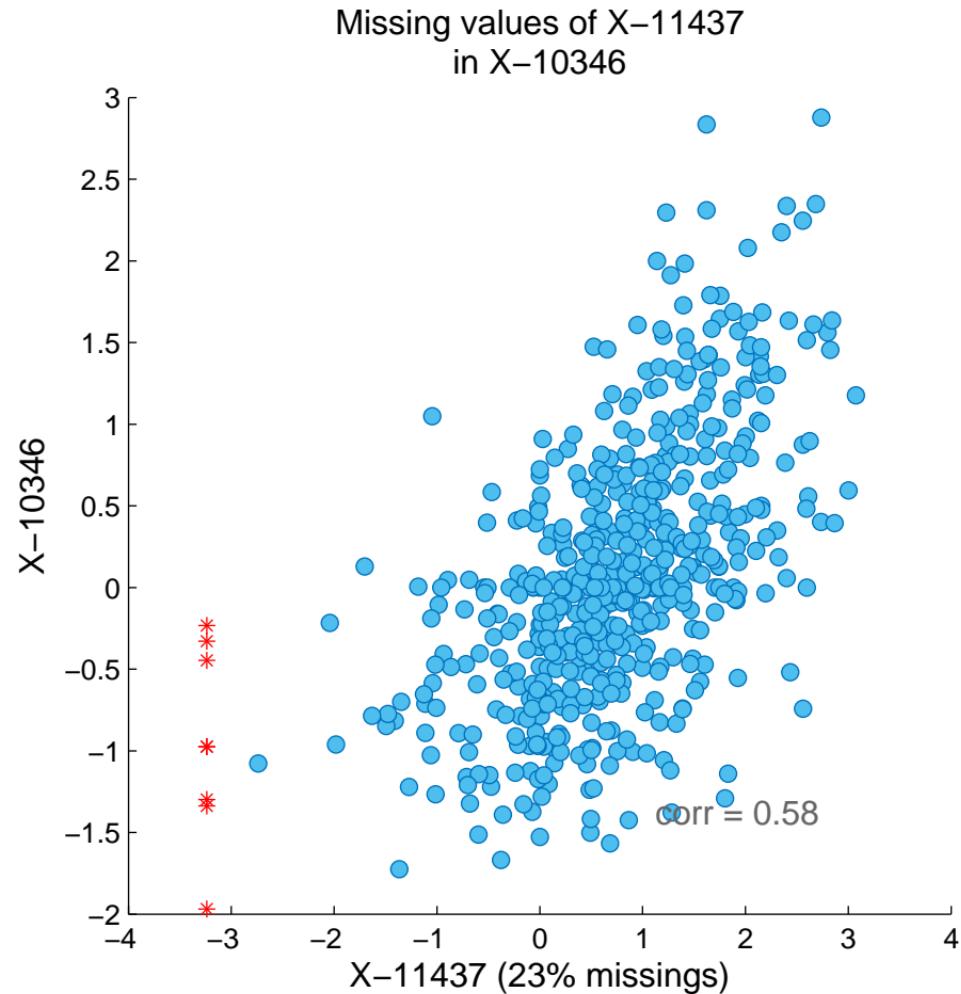
Missing values of 15-methylpalmitate (isobar with 2-methylpalmitate)
in myristate (14:0)

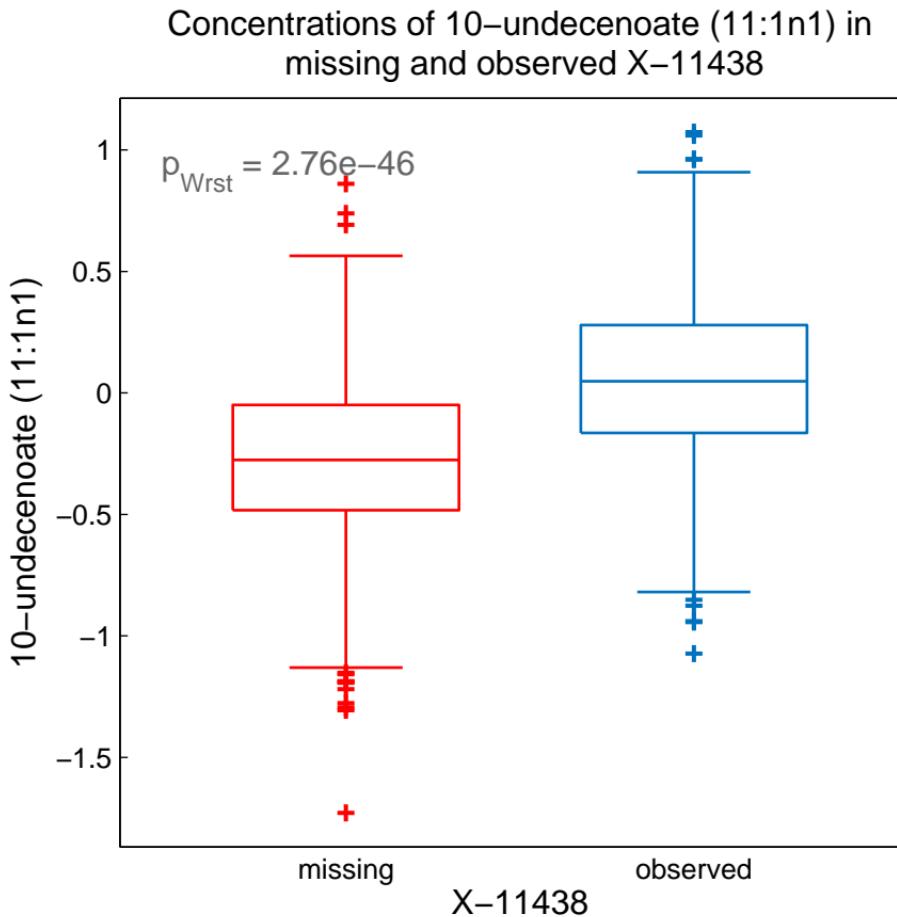
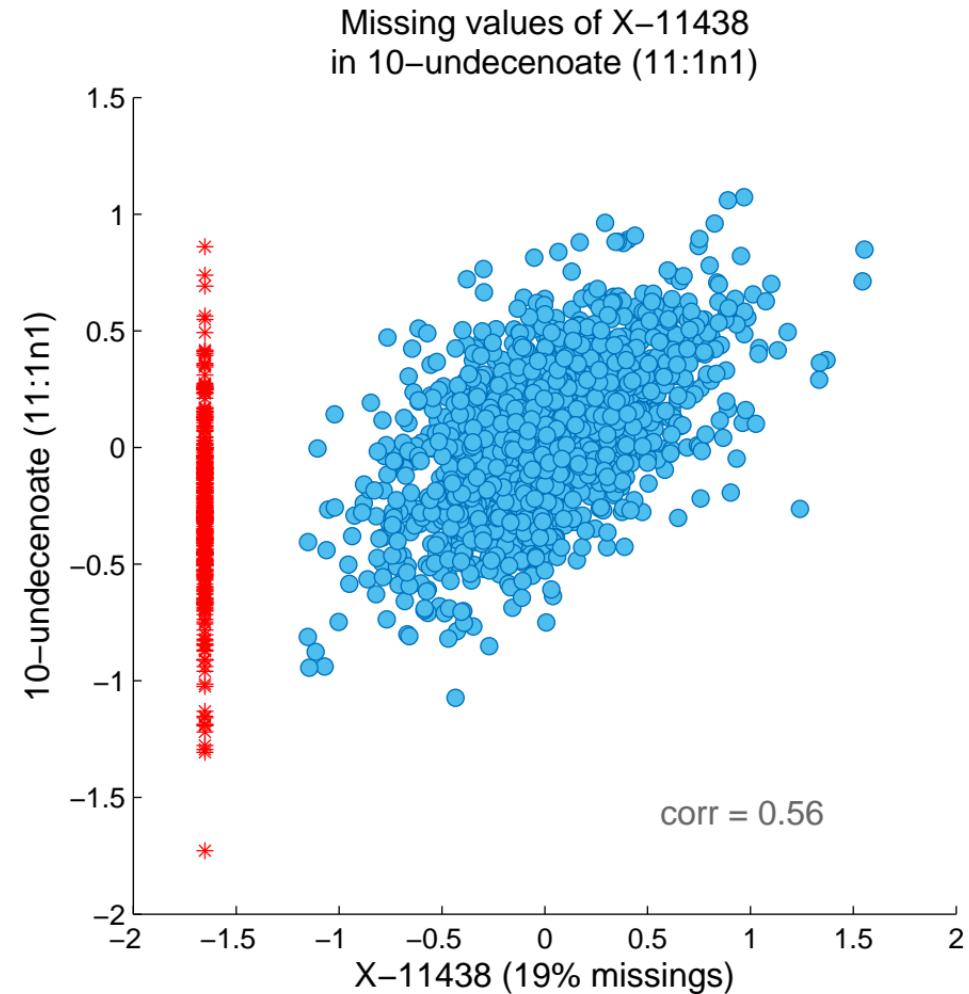


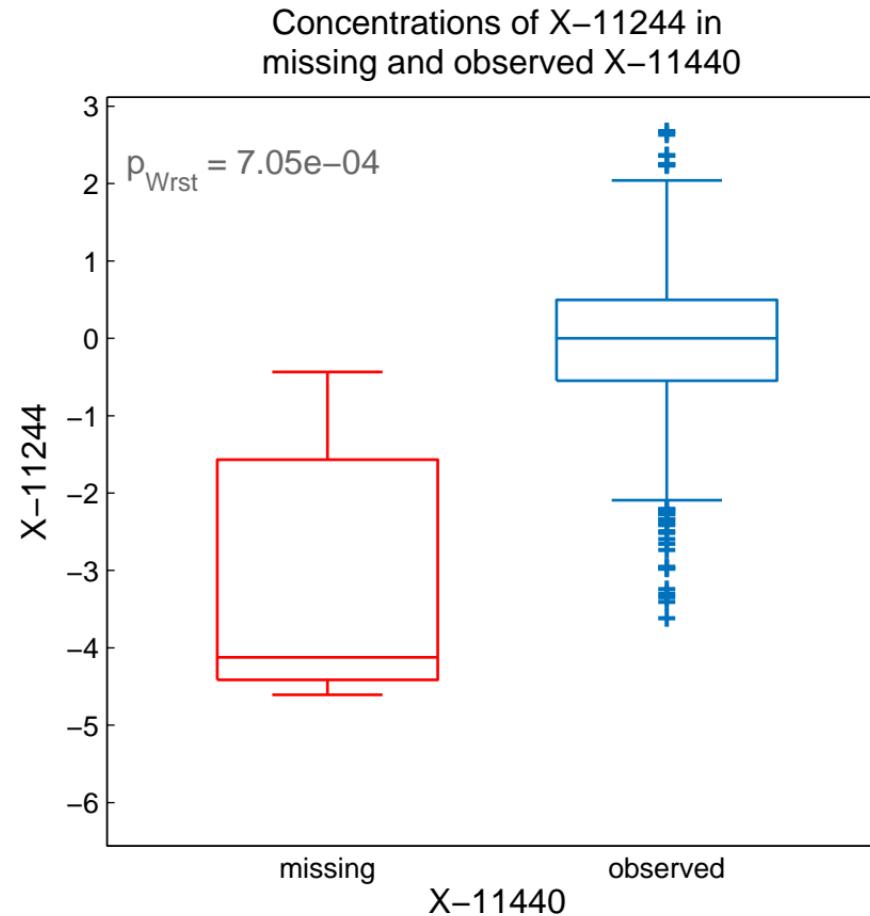
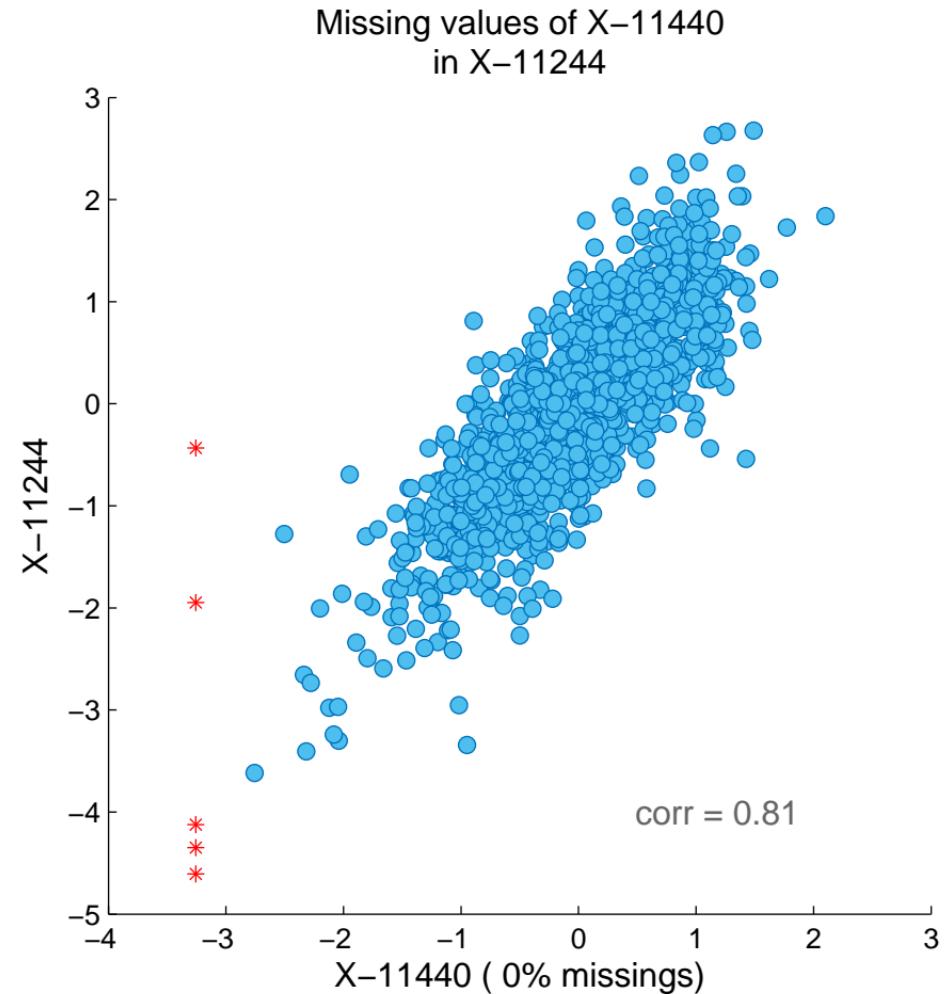
Concentrations of myristate (14:0) in
missing and observed 15-methylpalmitate (isobar with 2-methylpalmitate)

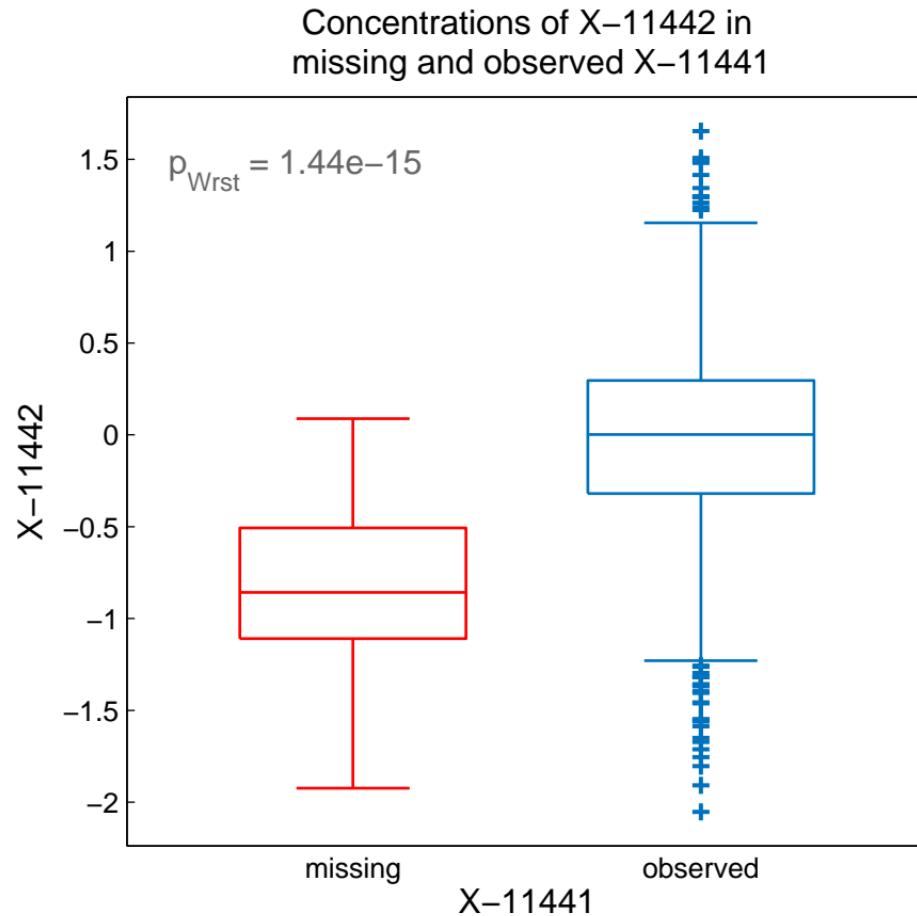
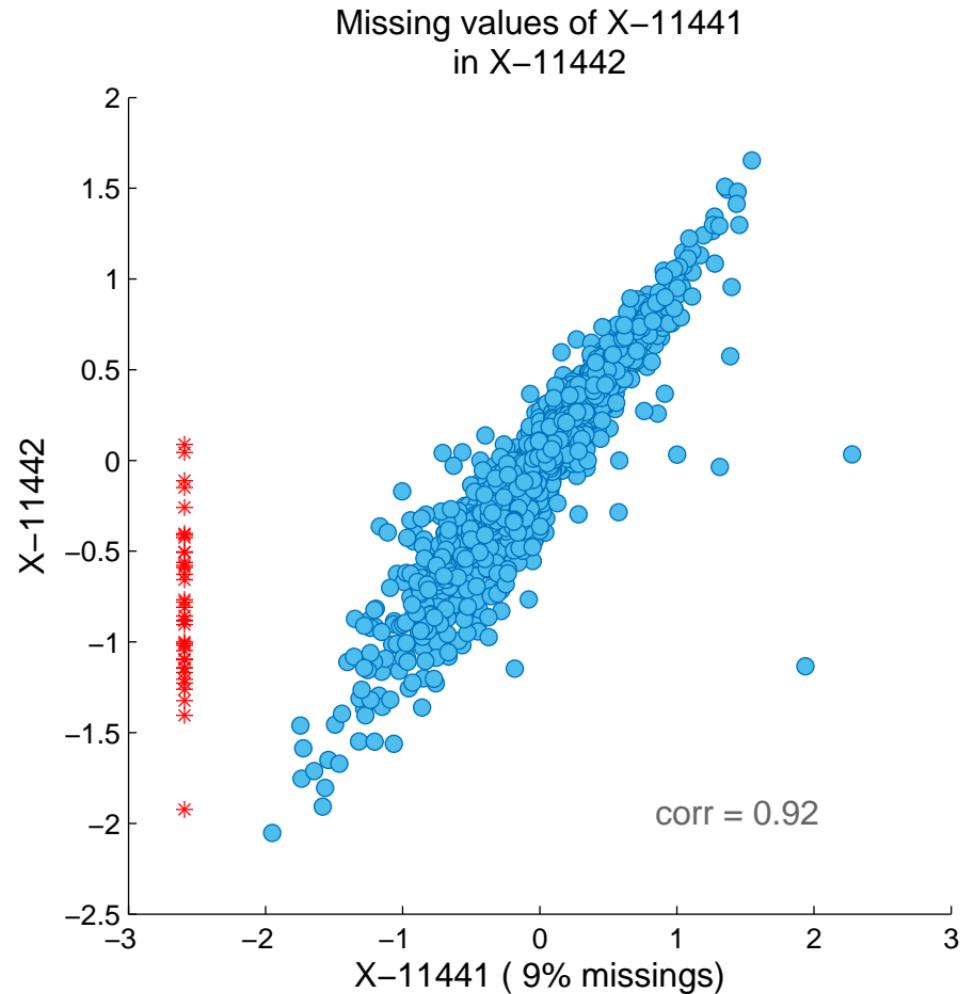


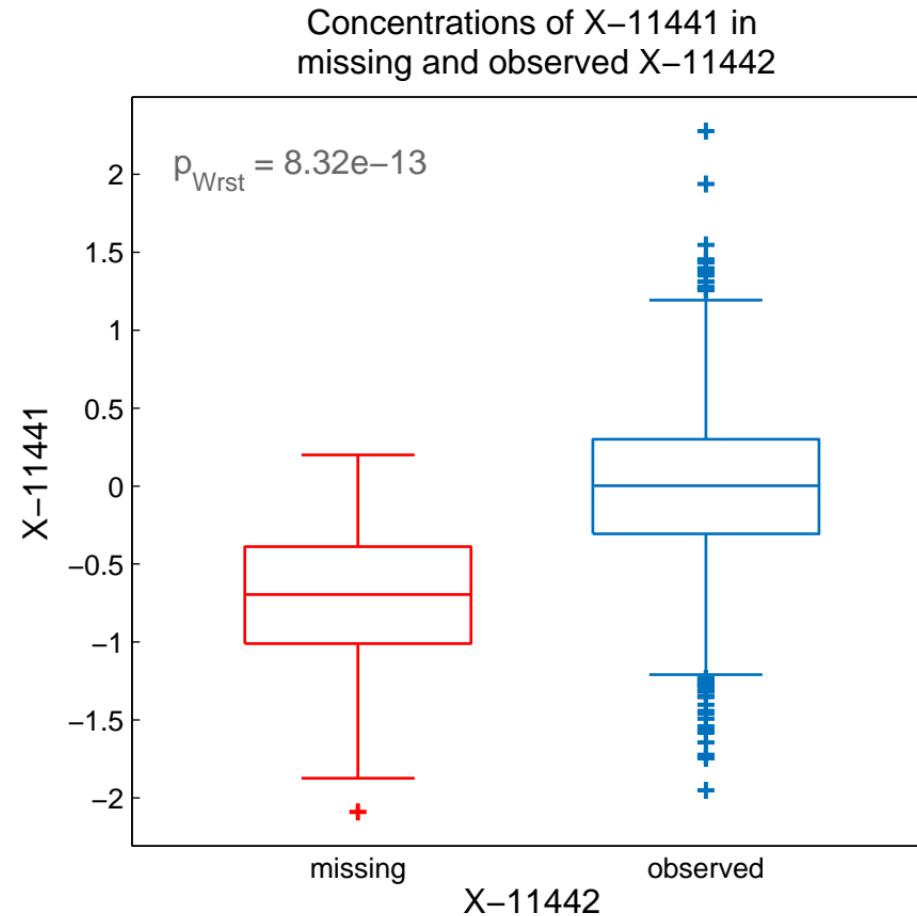
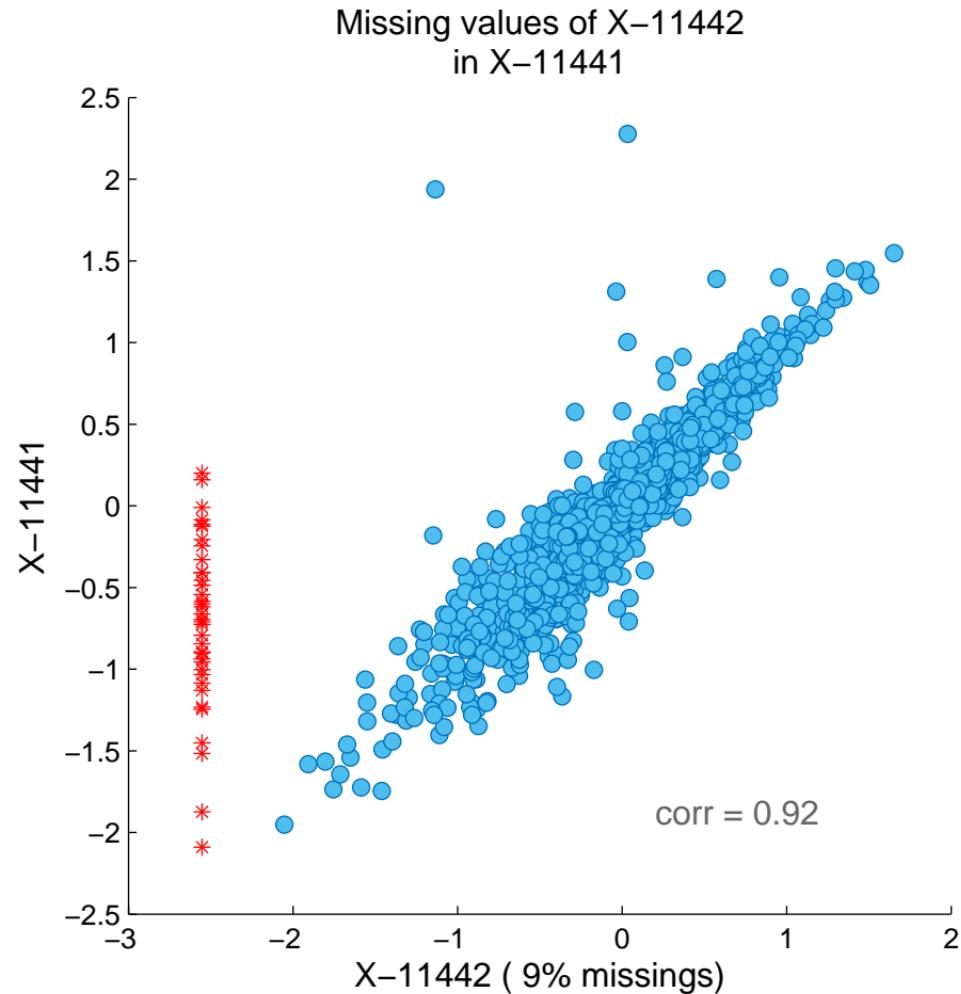


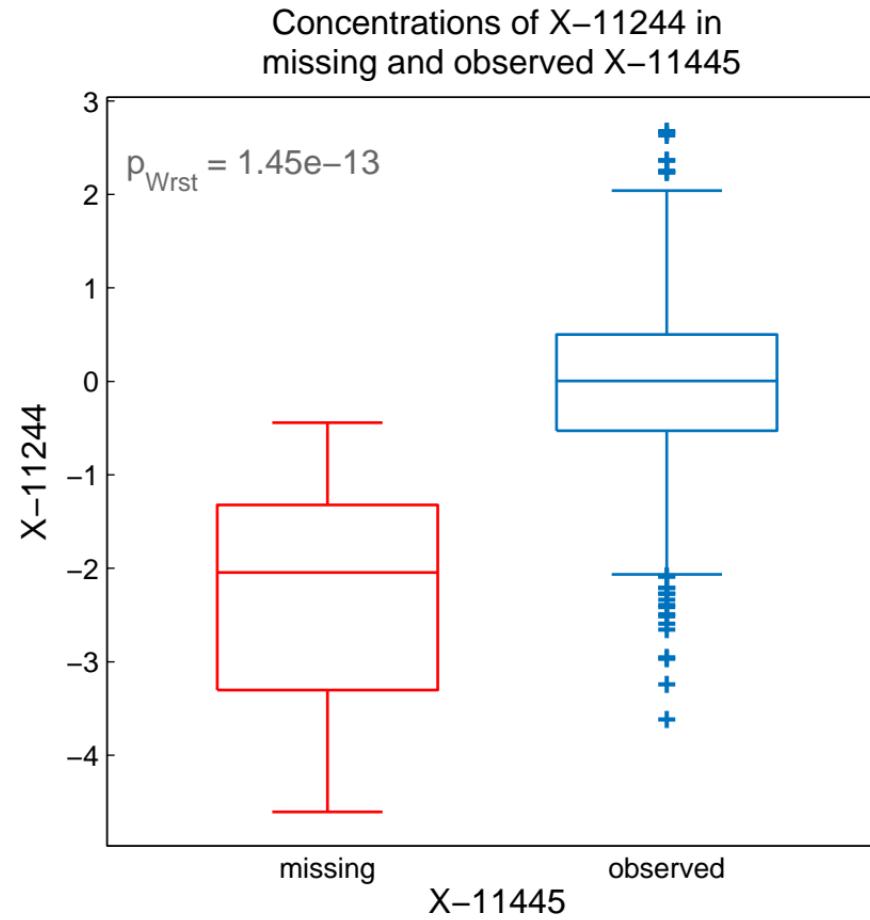
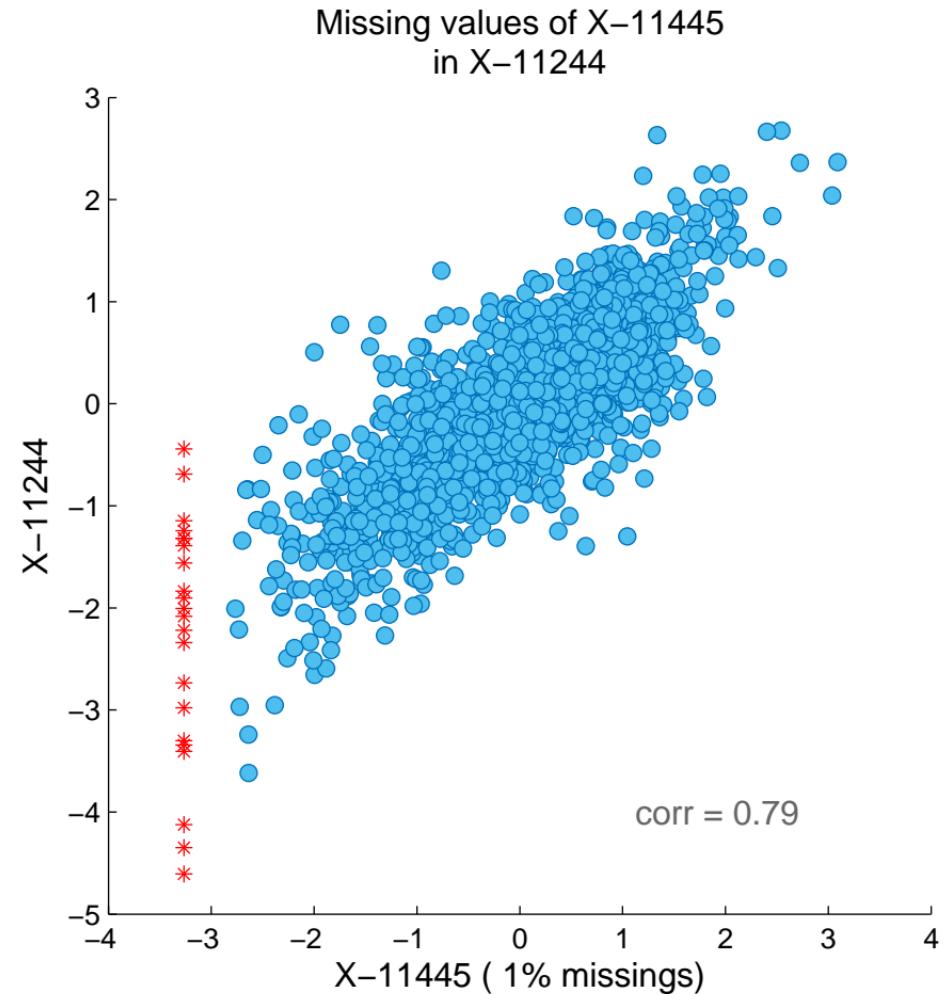


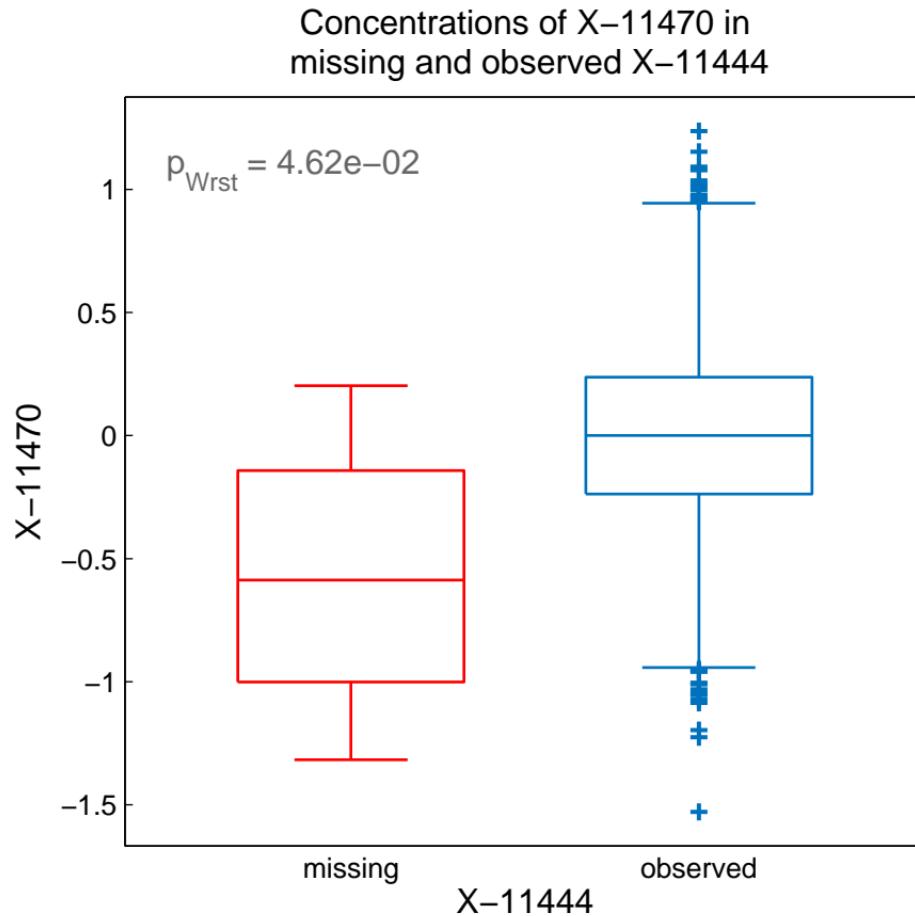
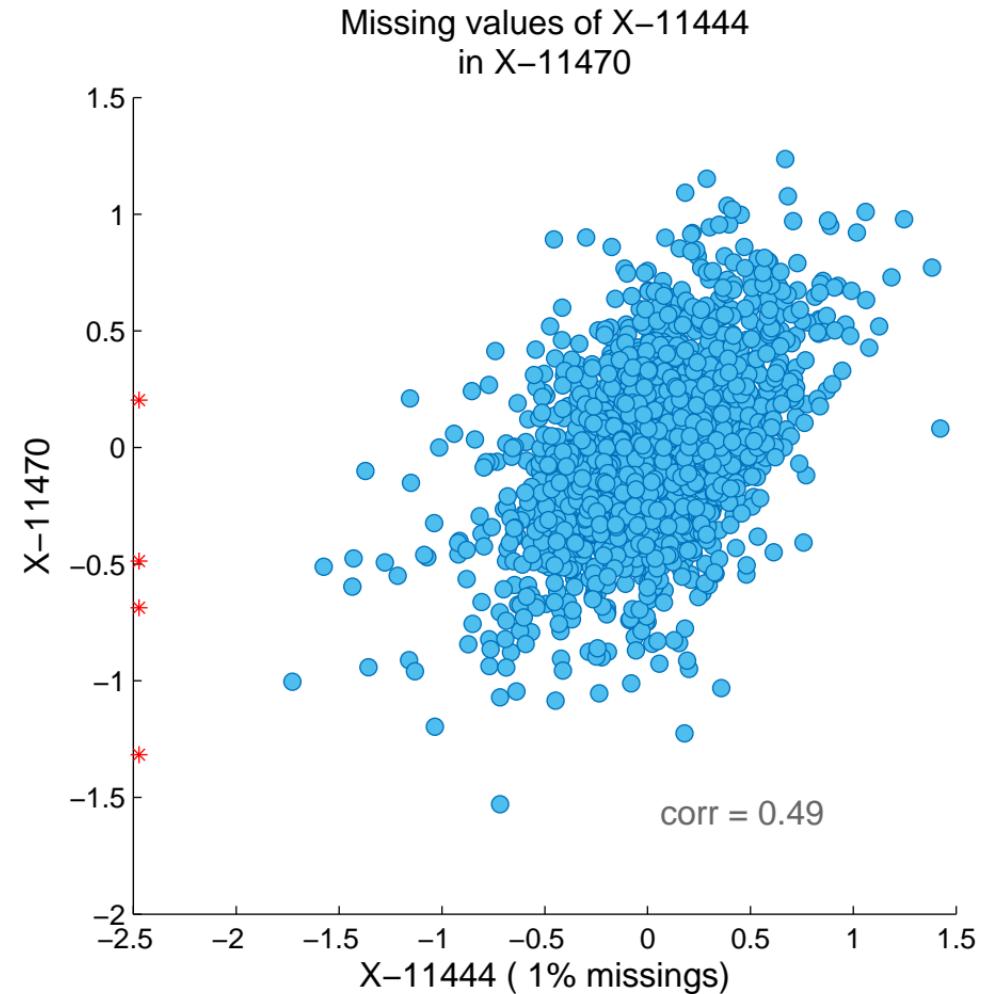


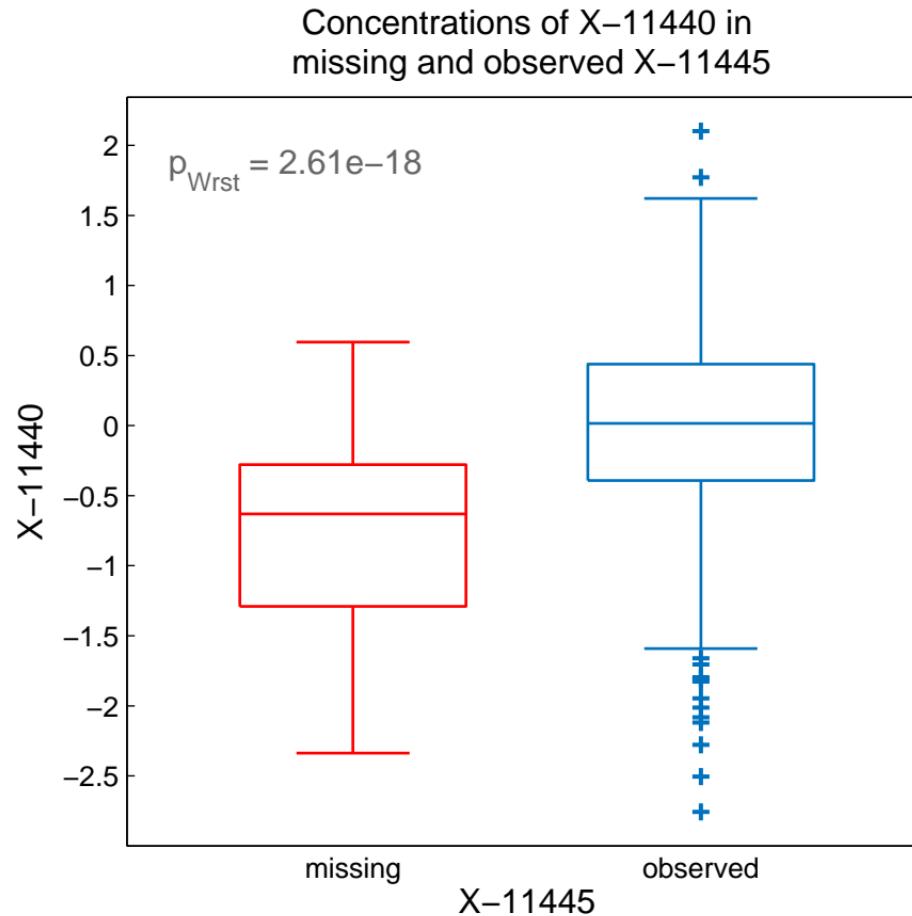
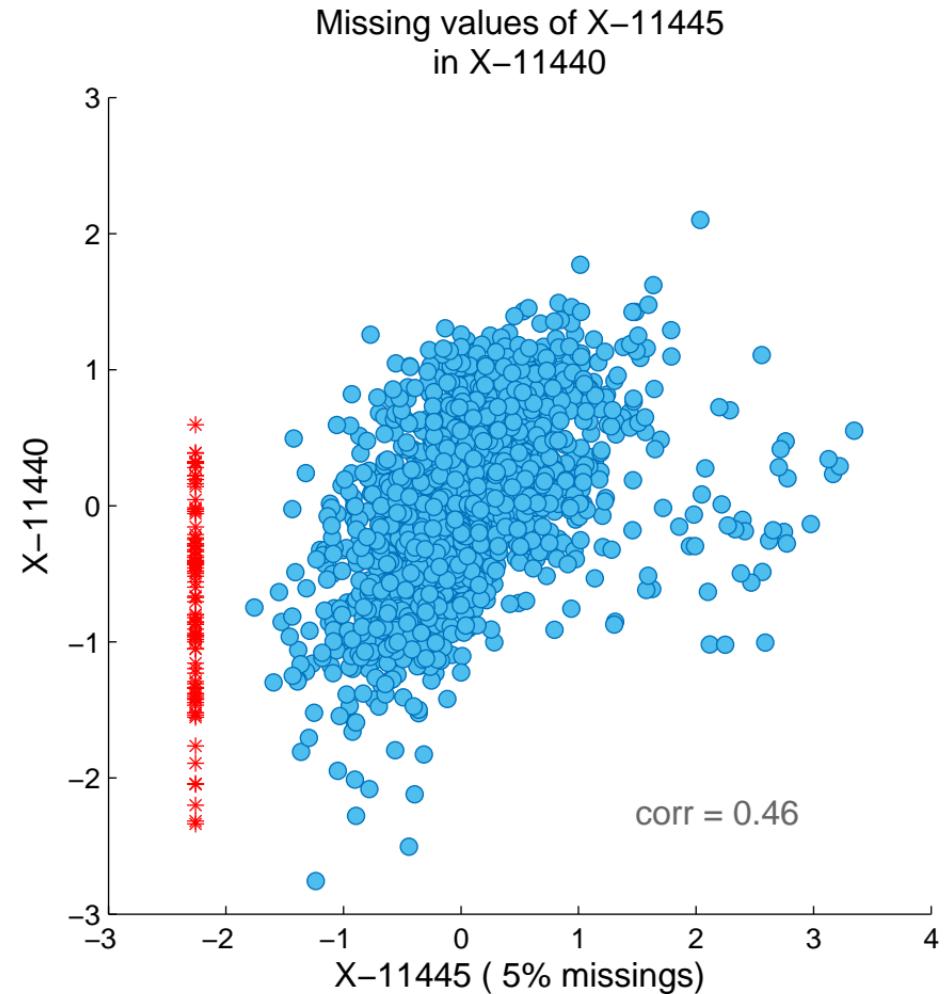


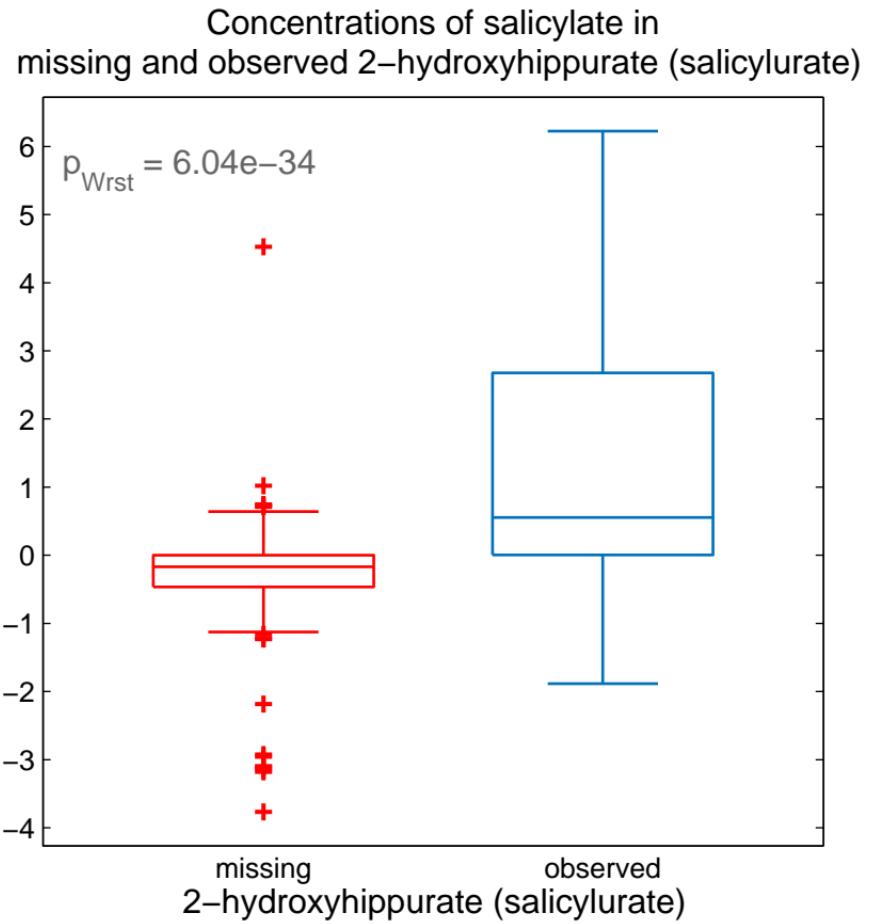
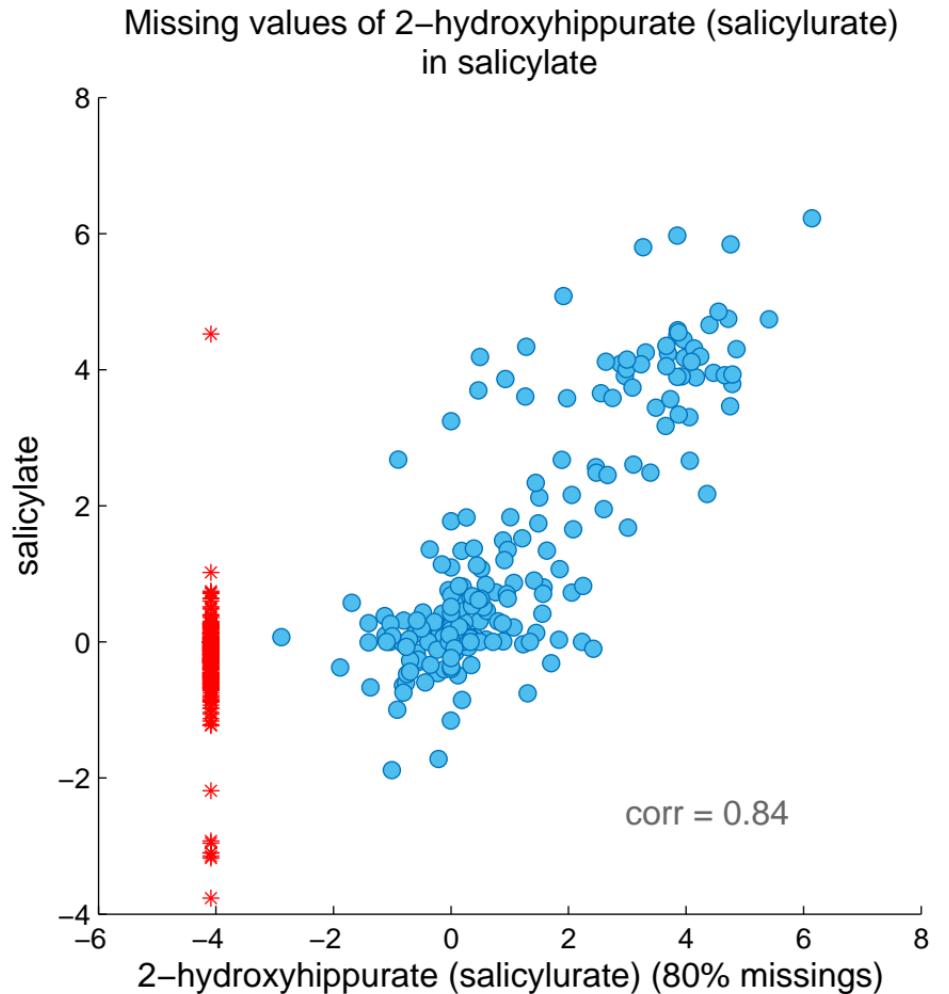


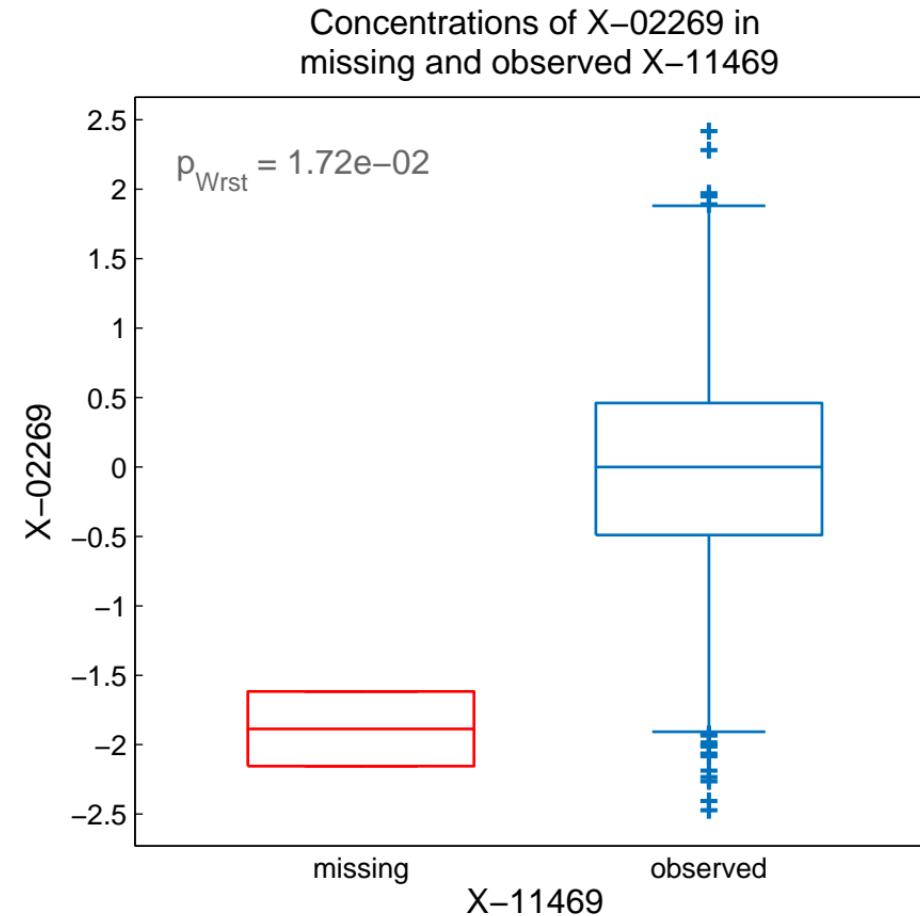
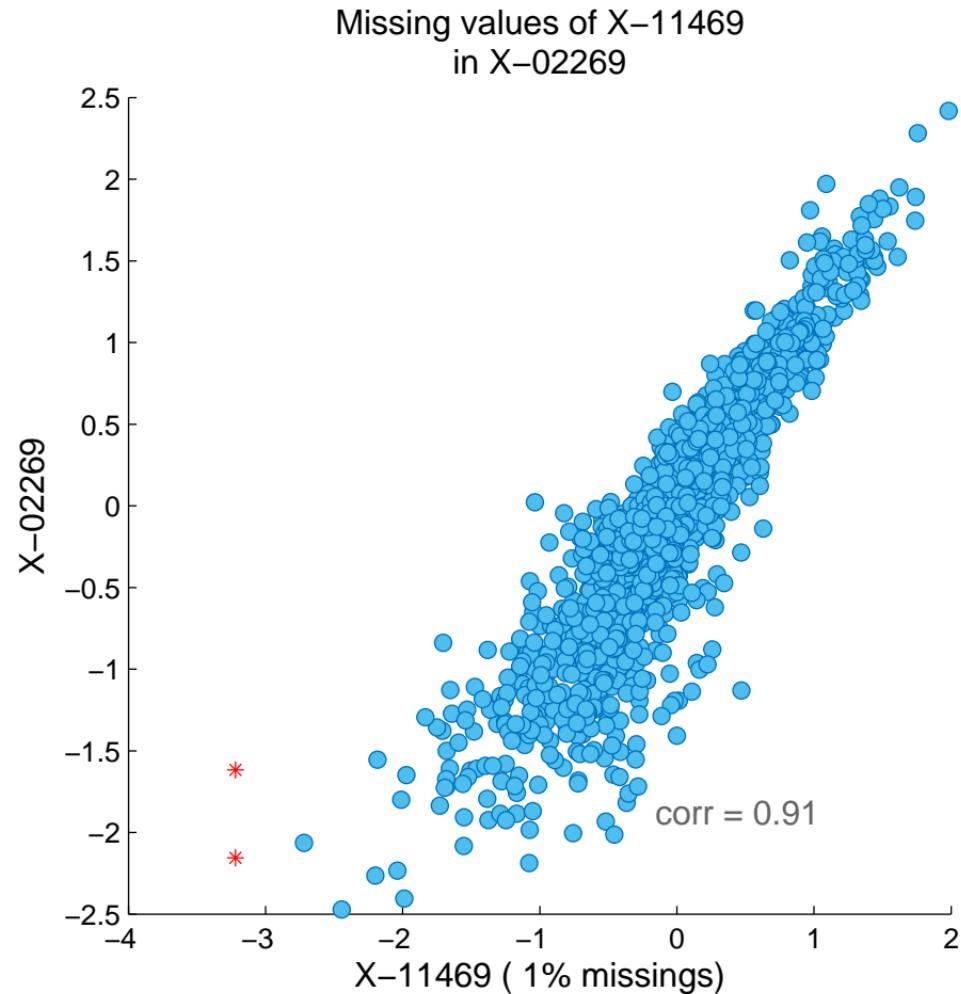


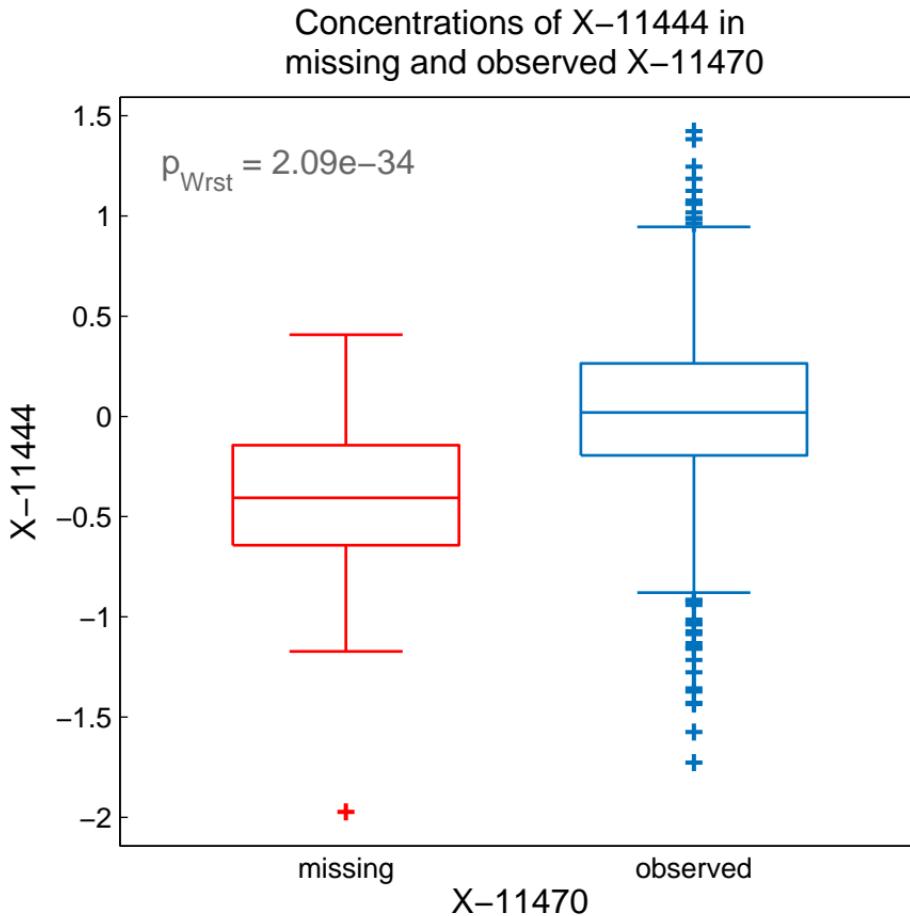
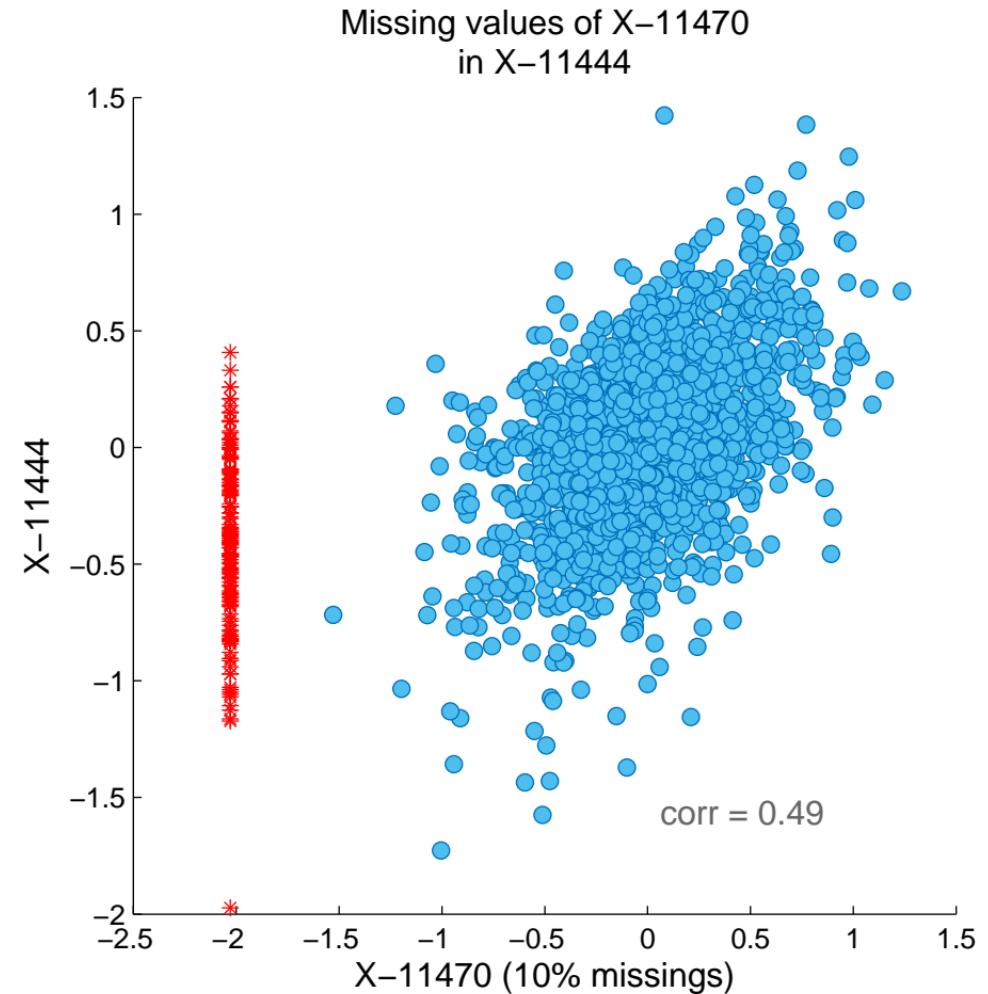


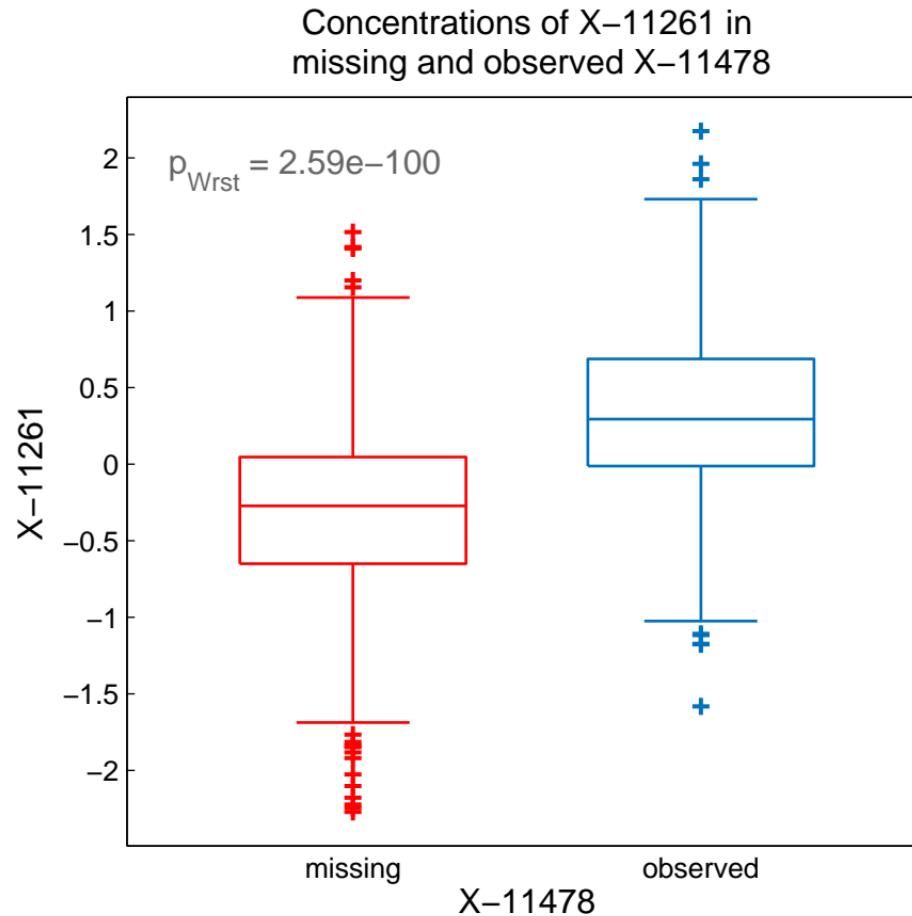
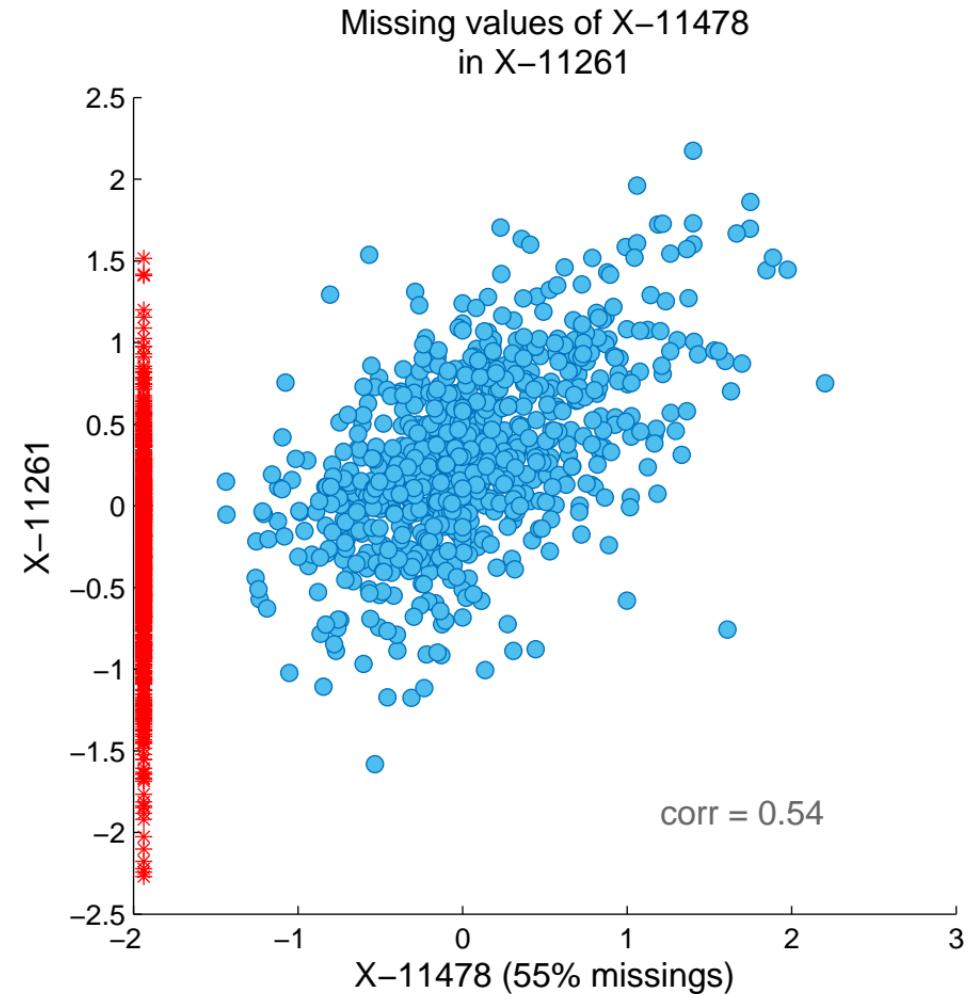


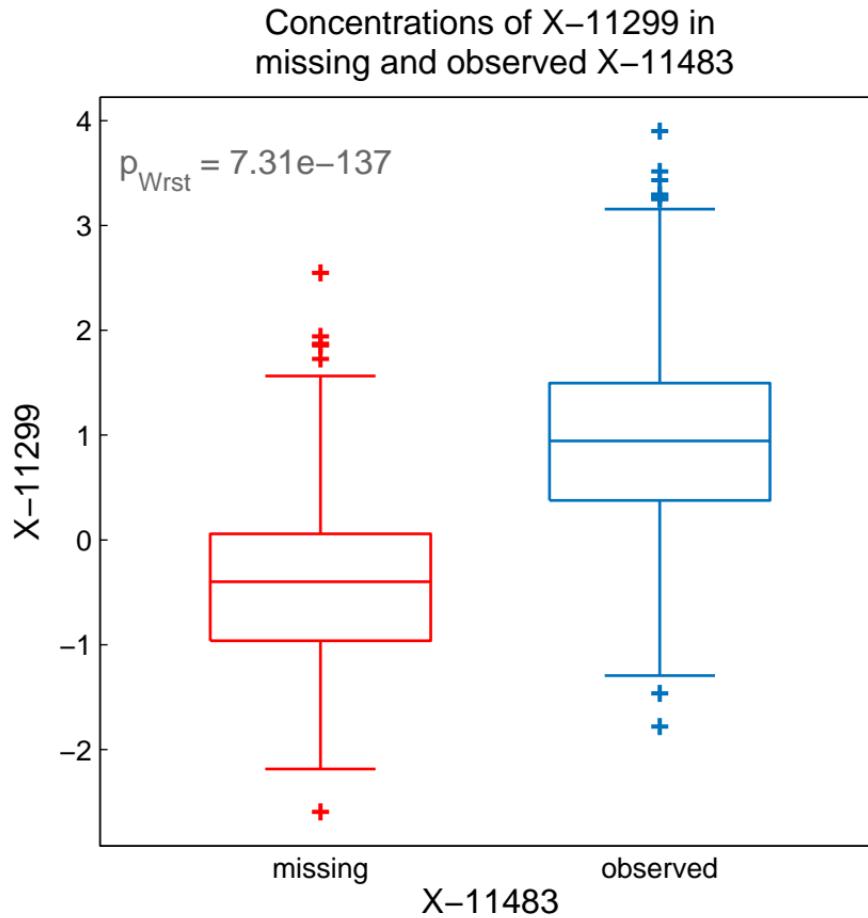
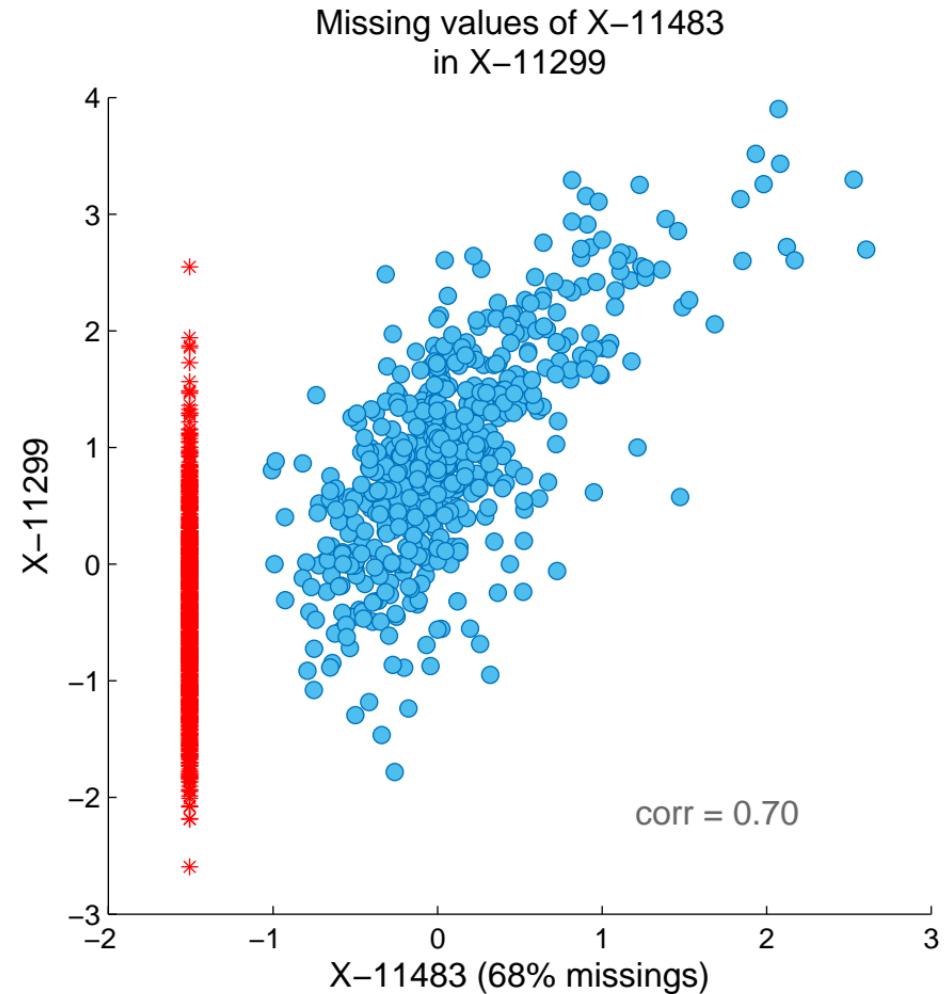


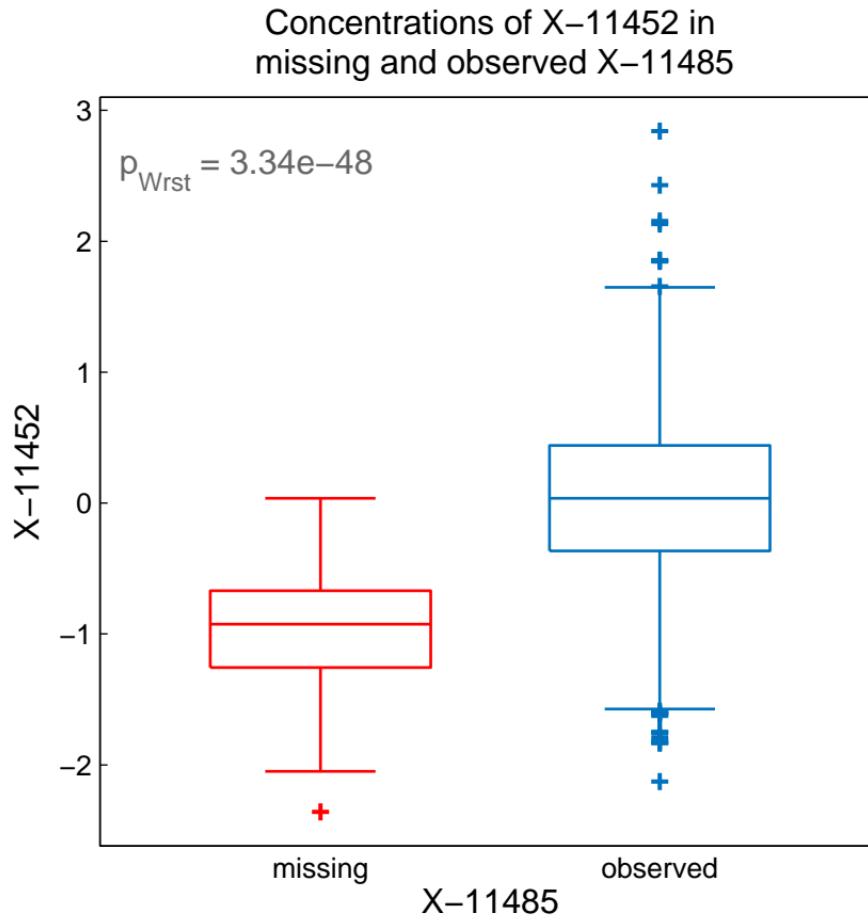
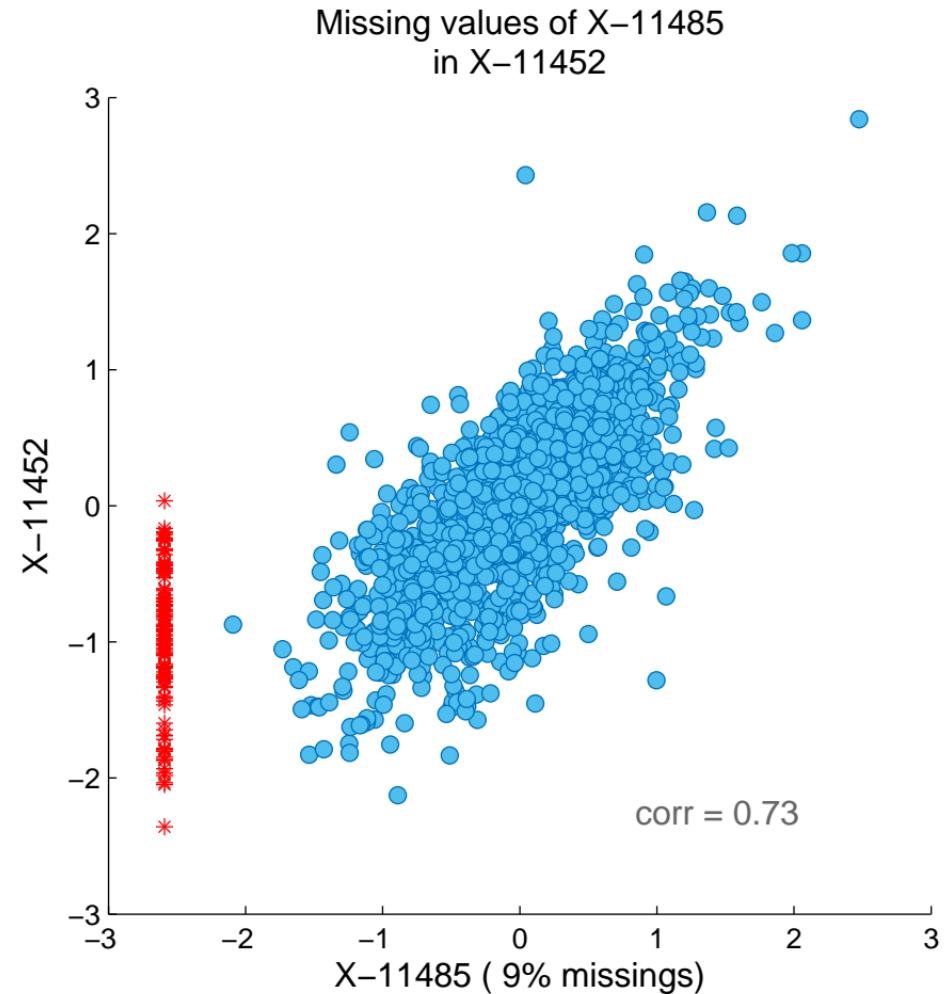


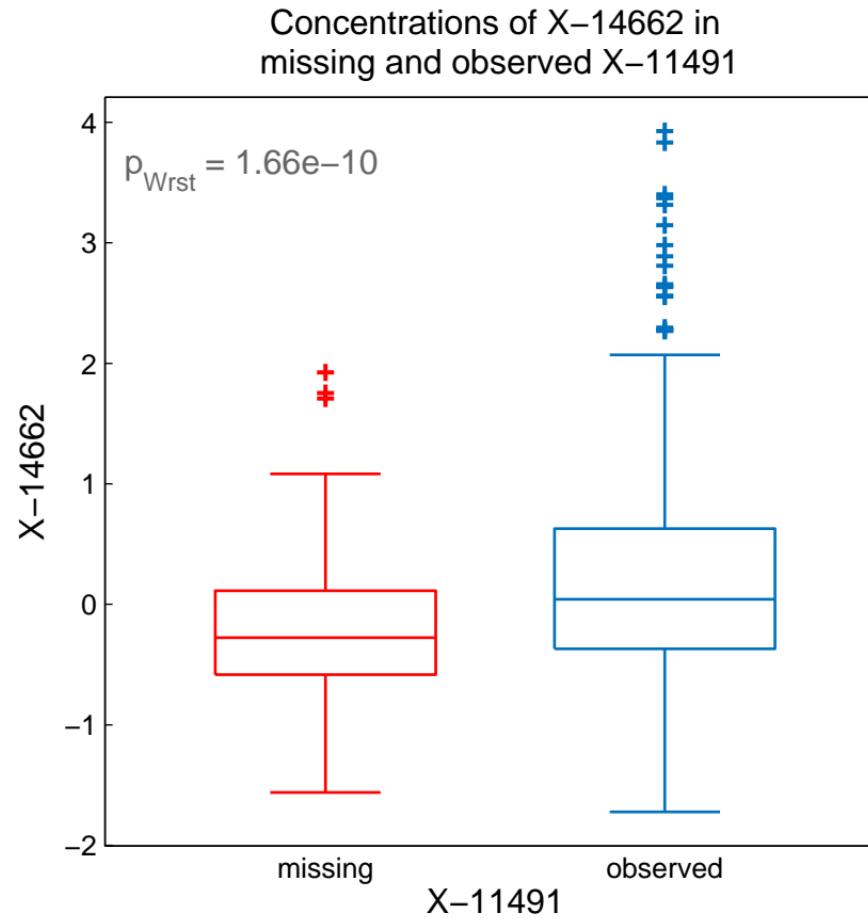
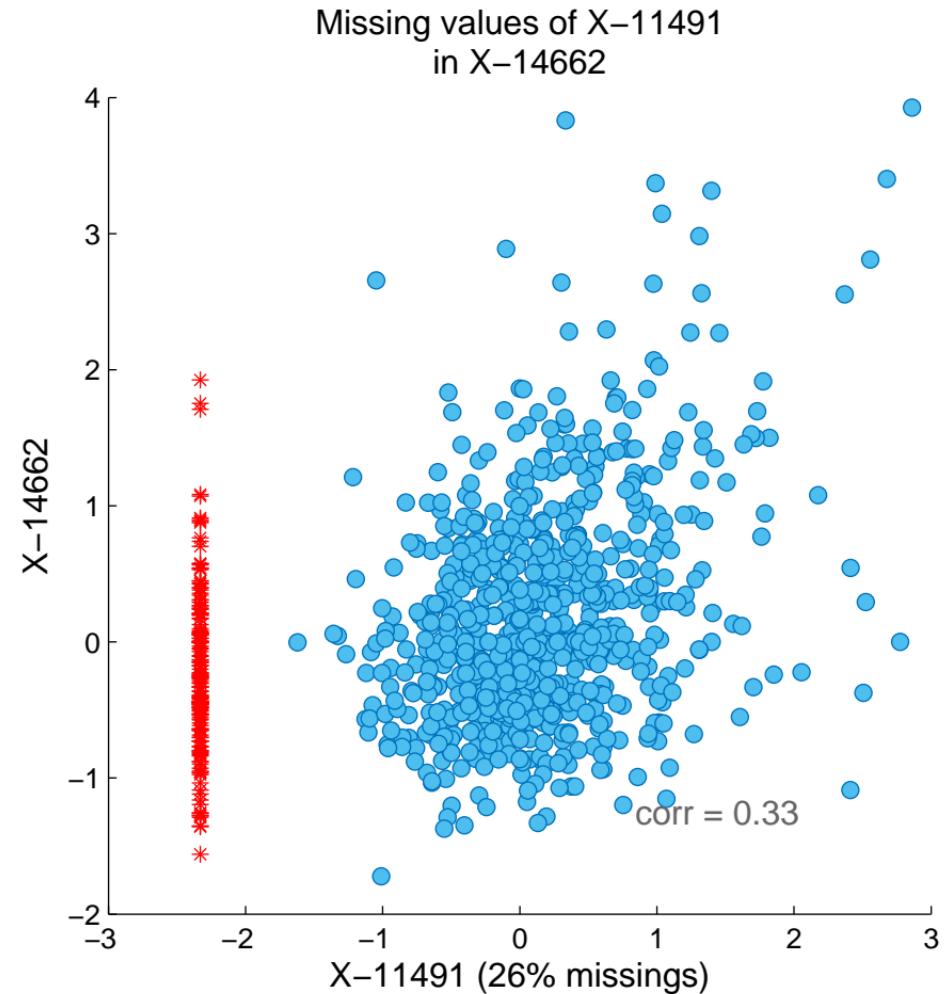


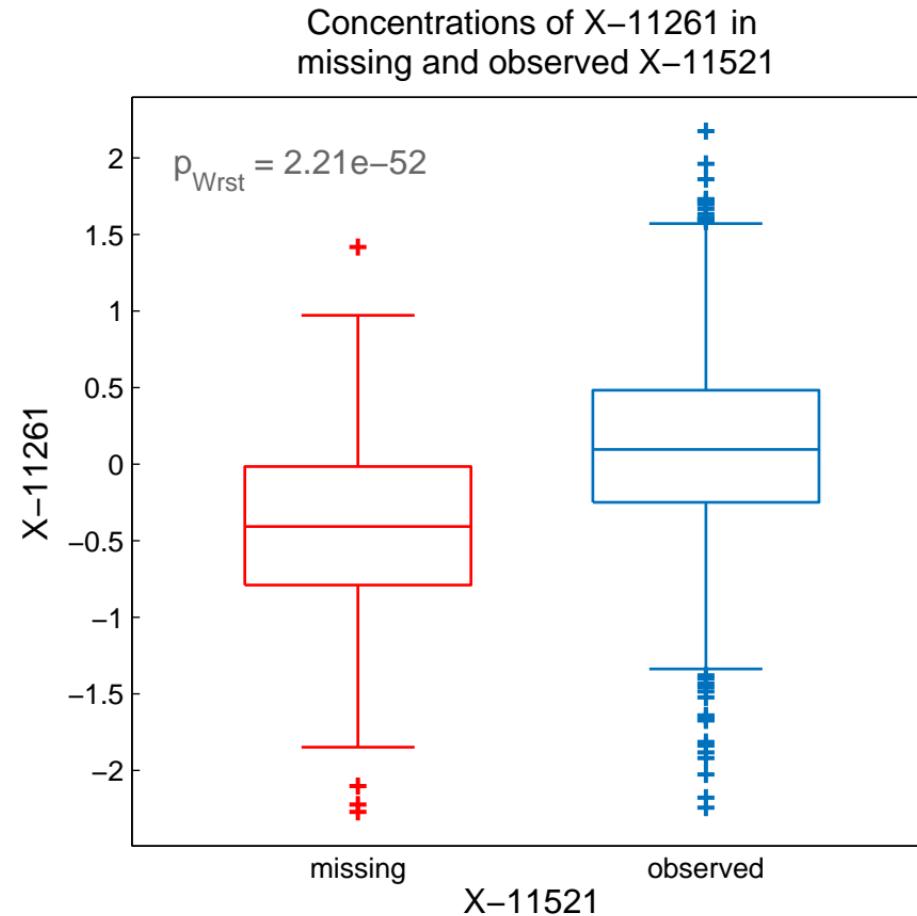
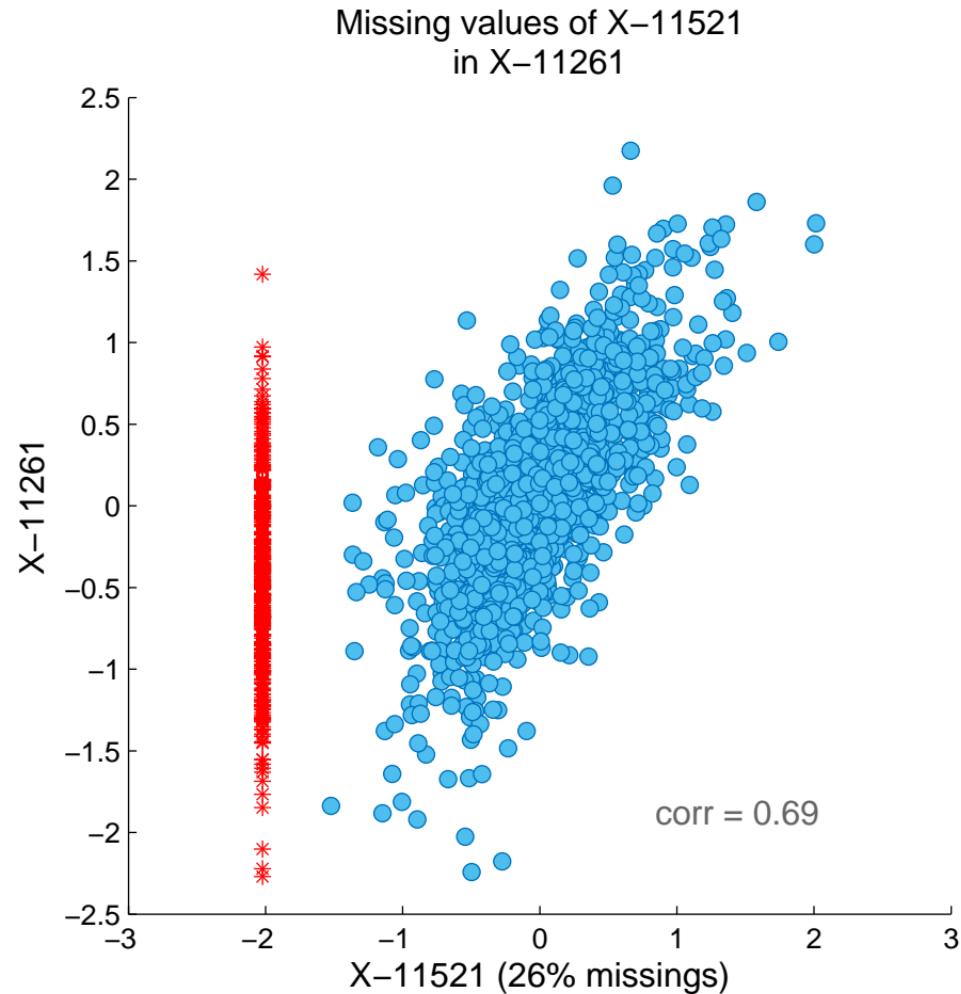


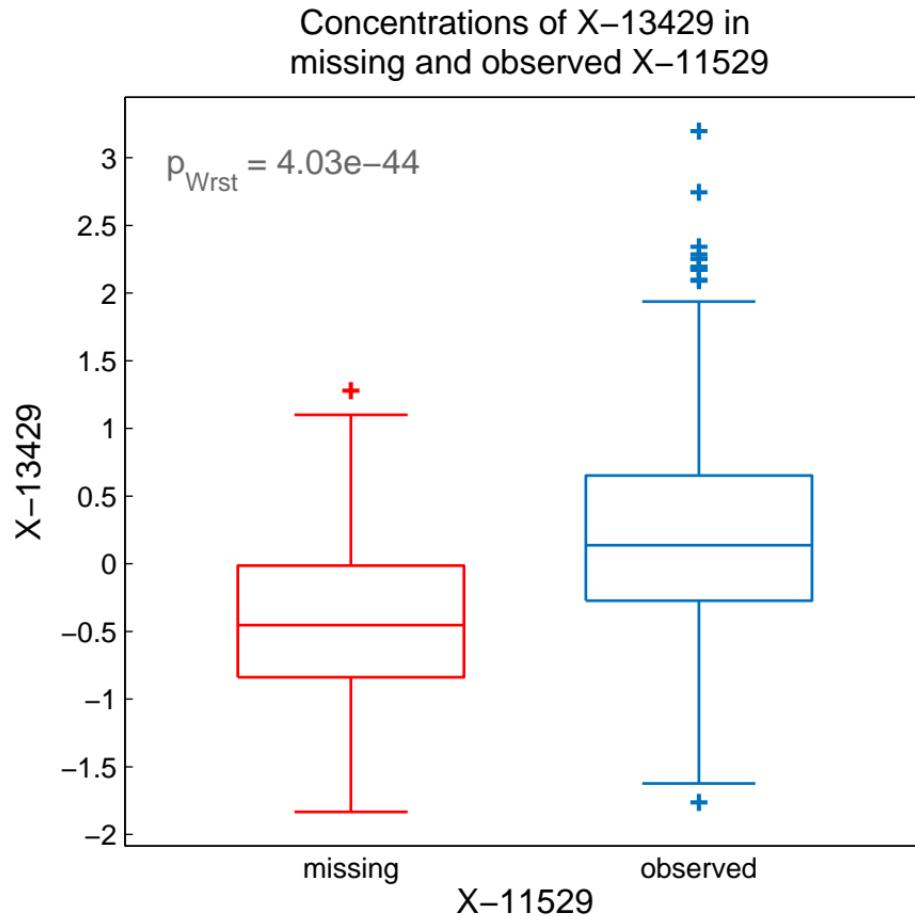
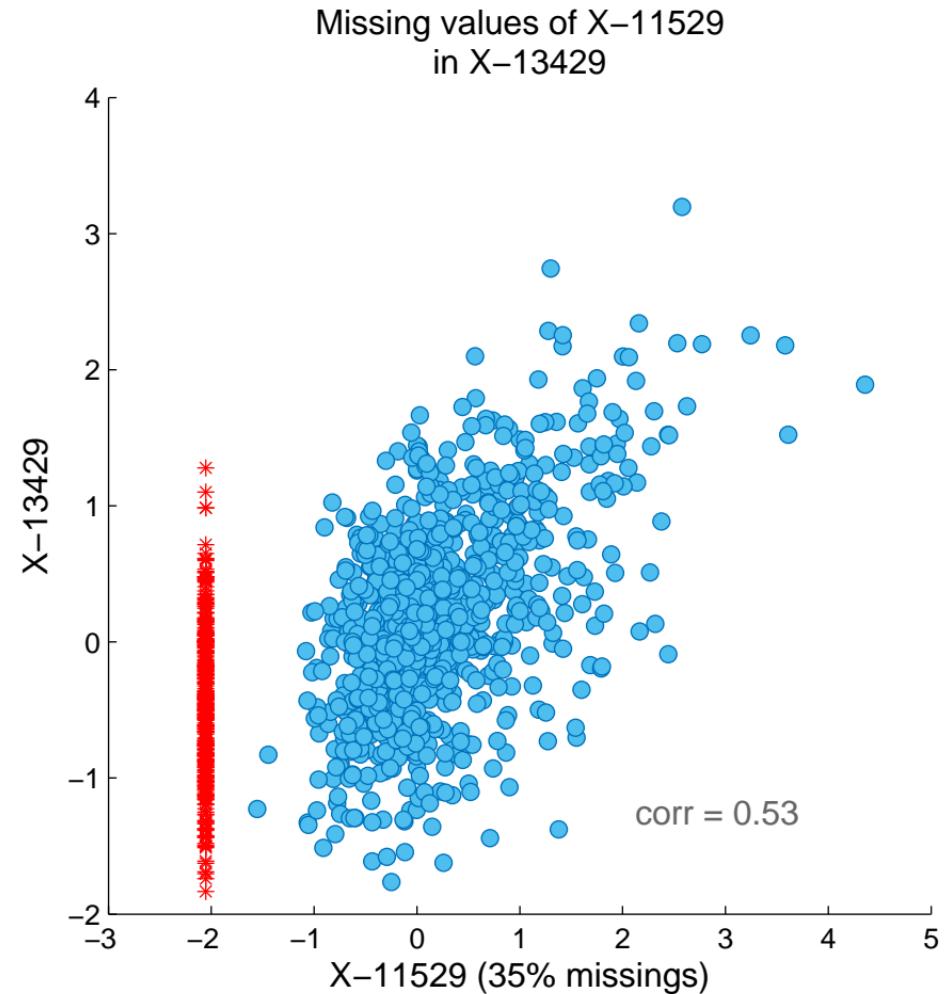


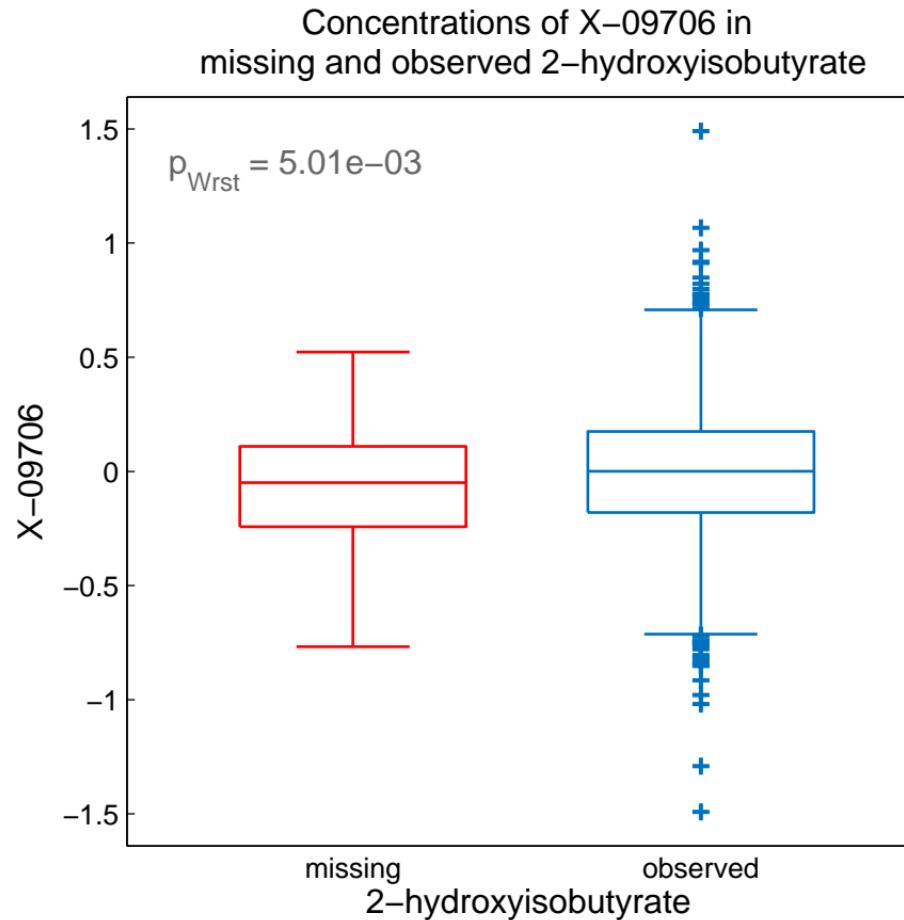
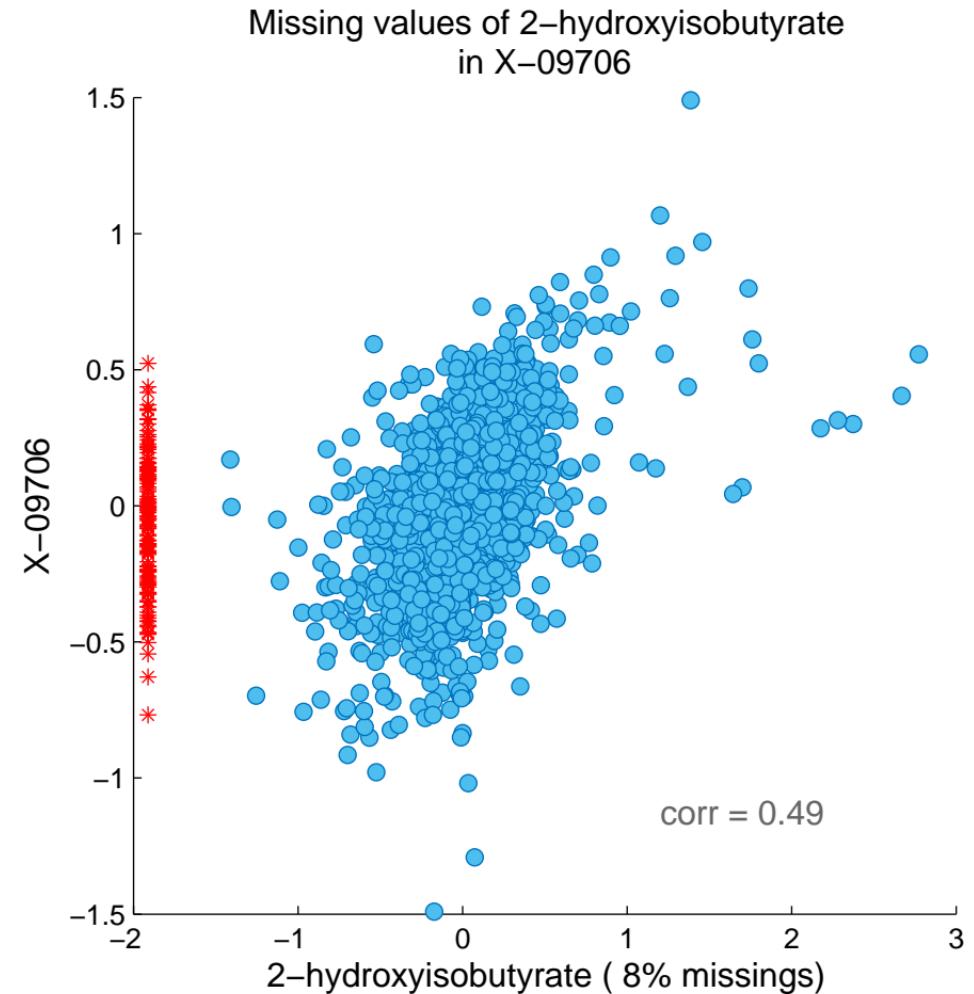


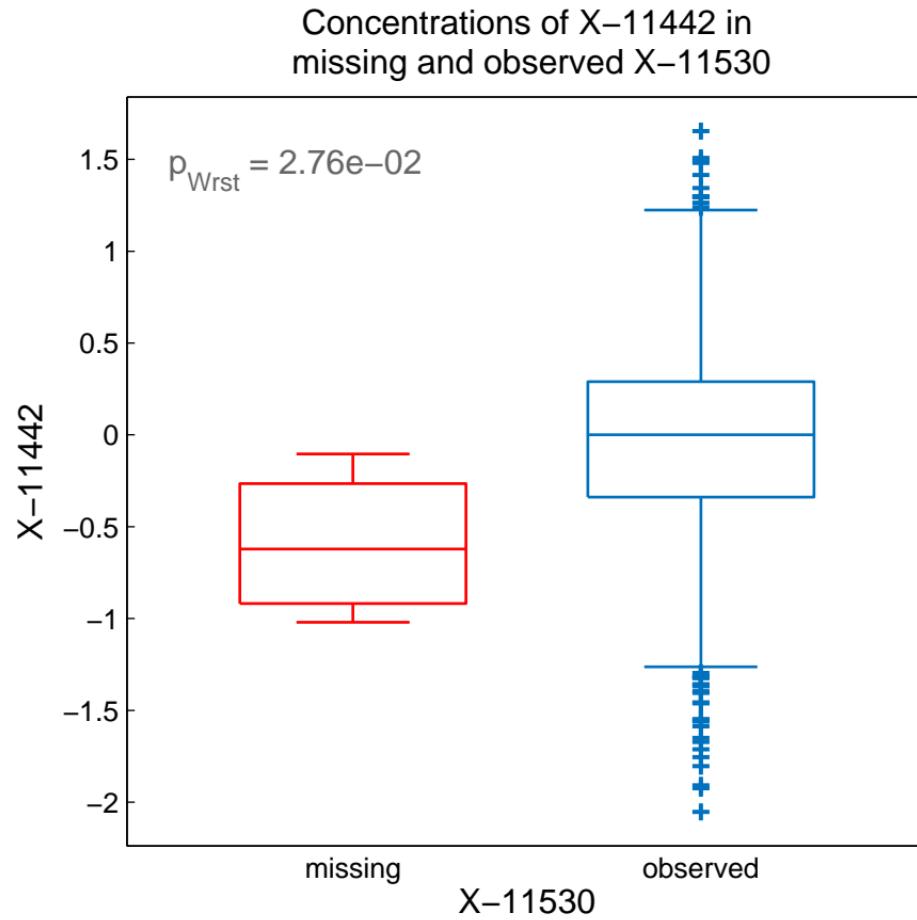
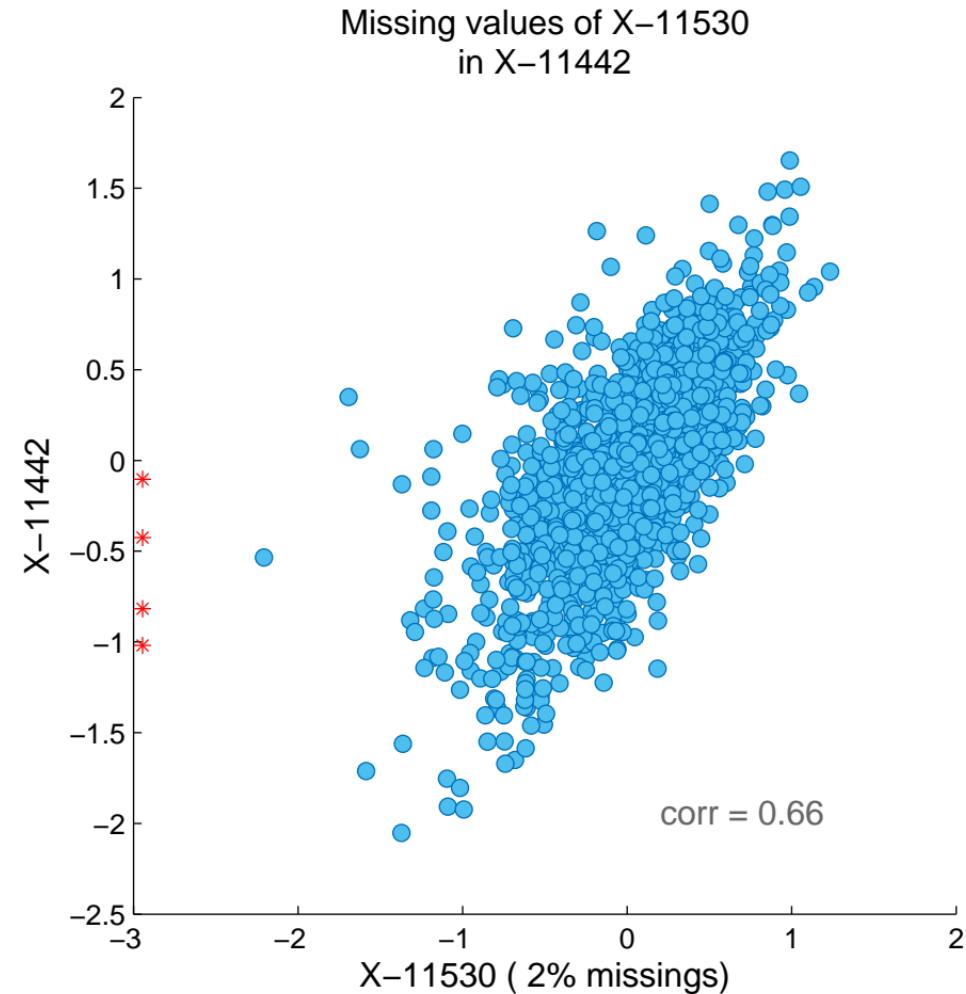


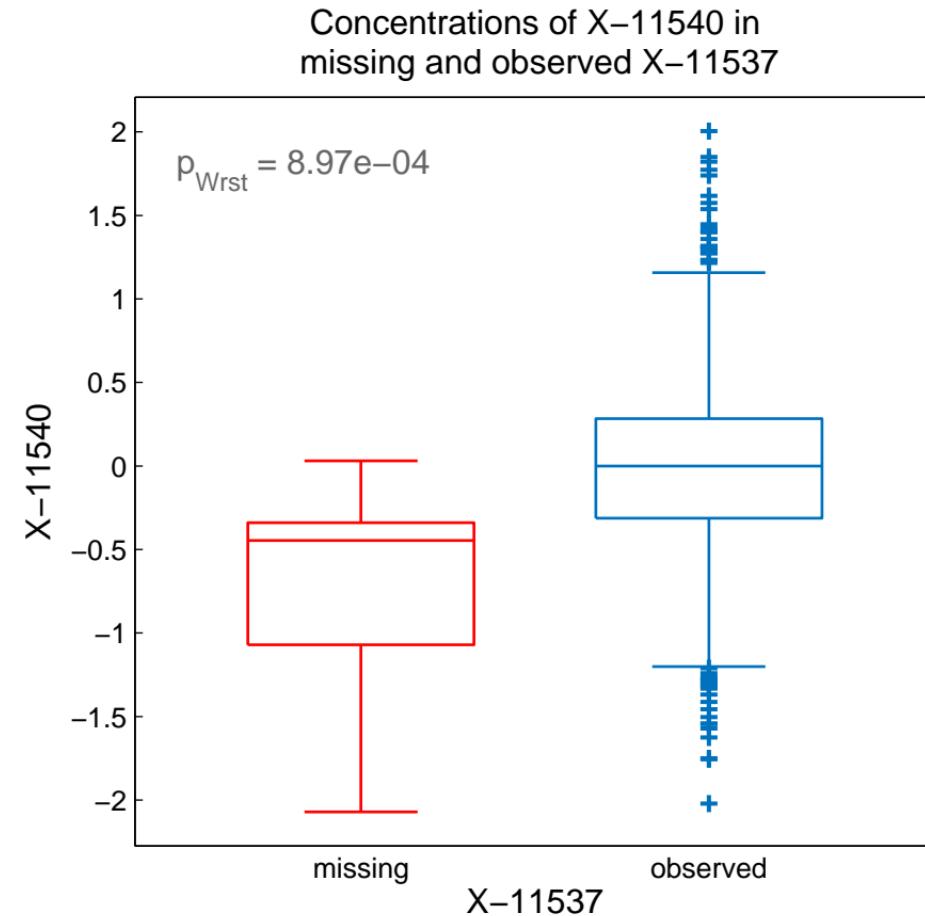
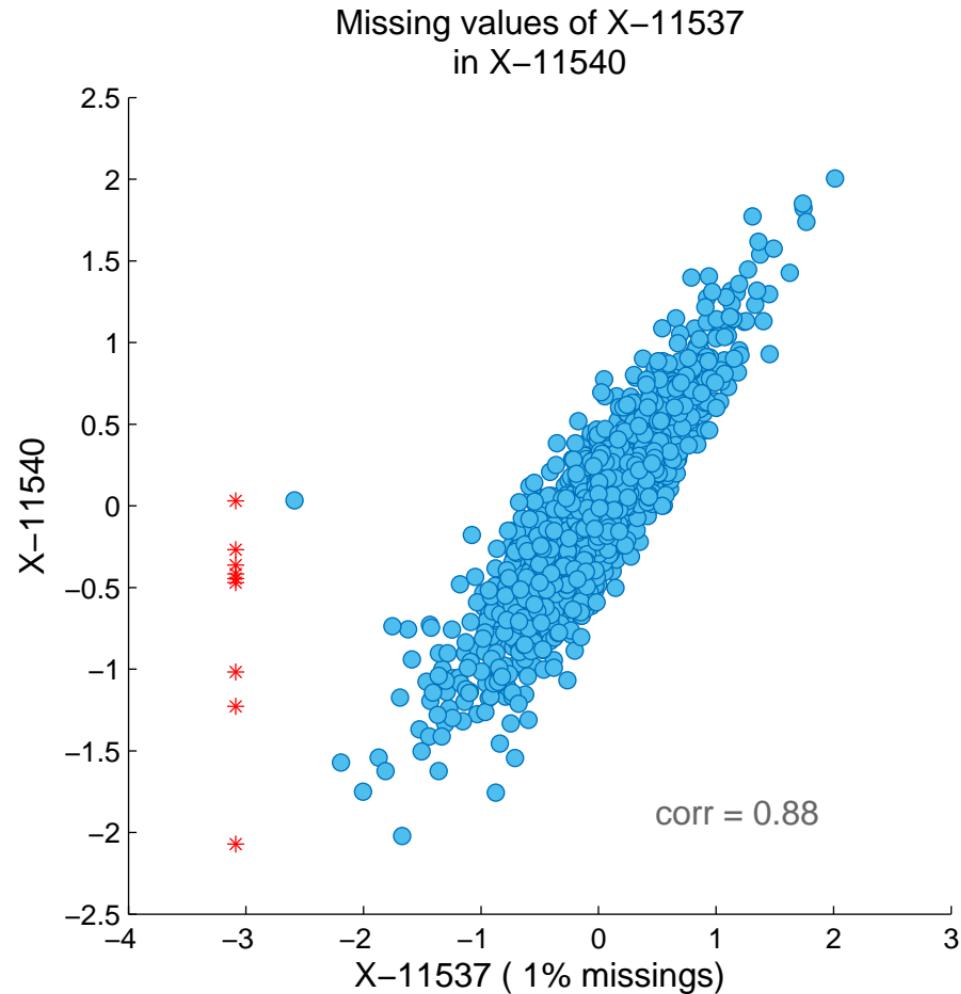


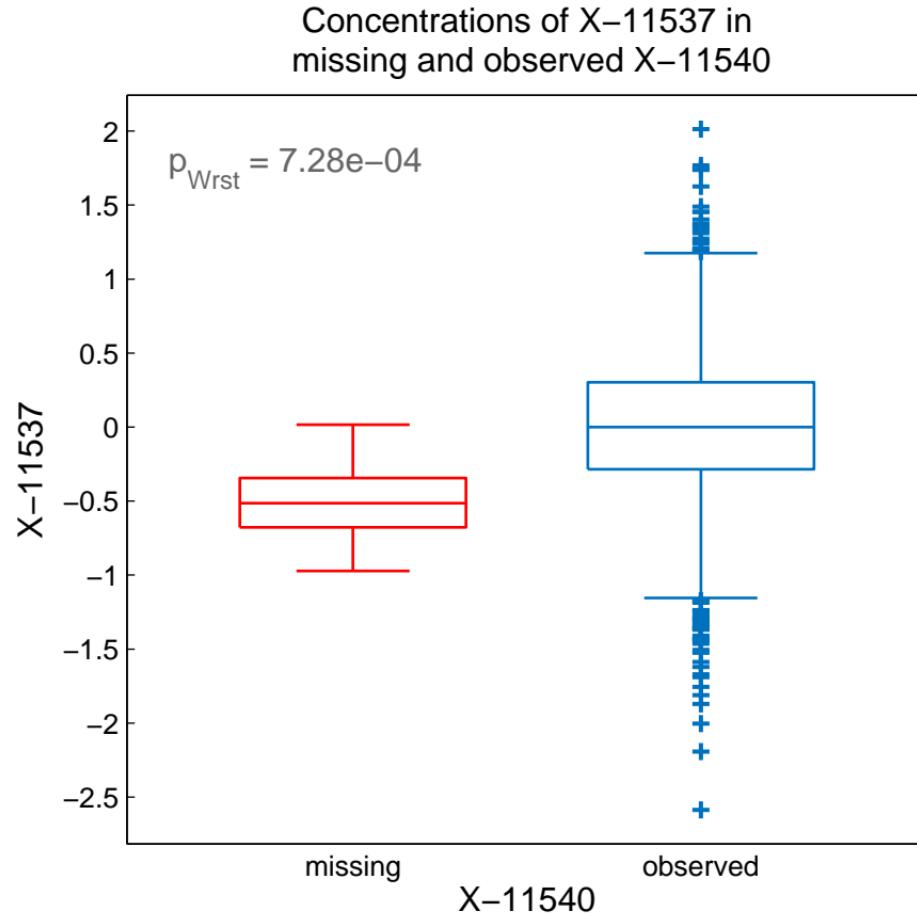
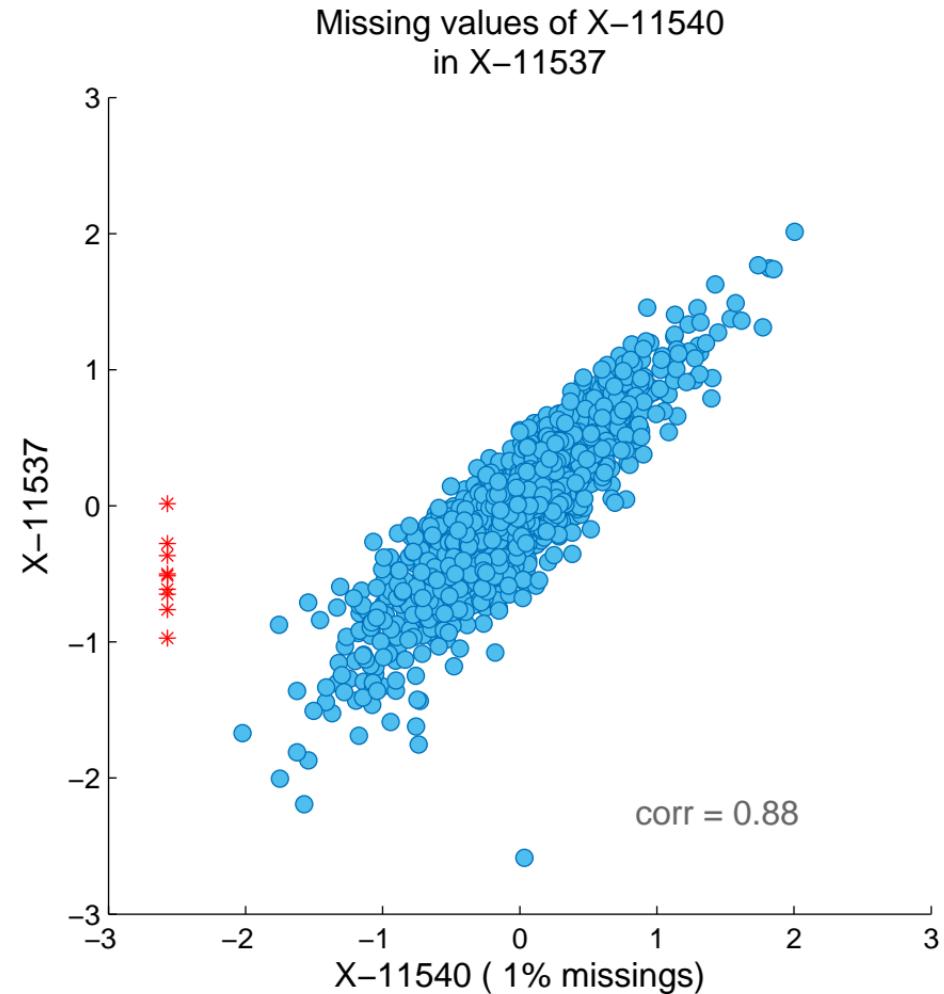


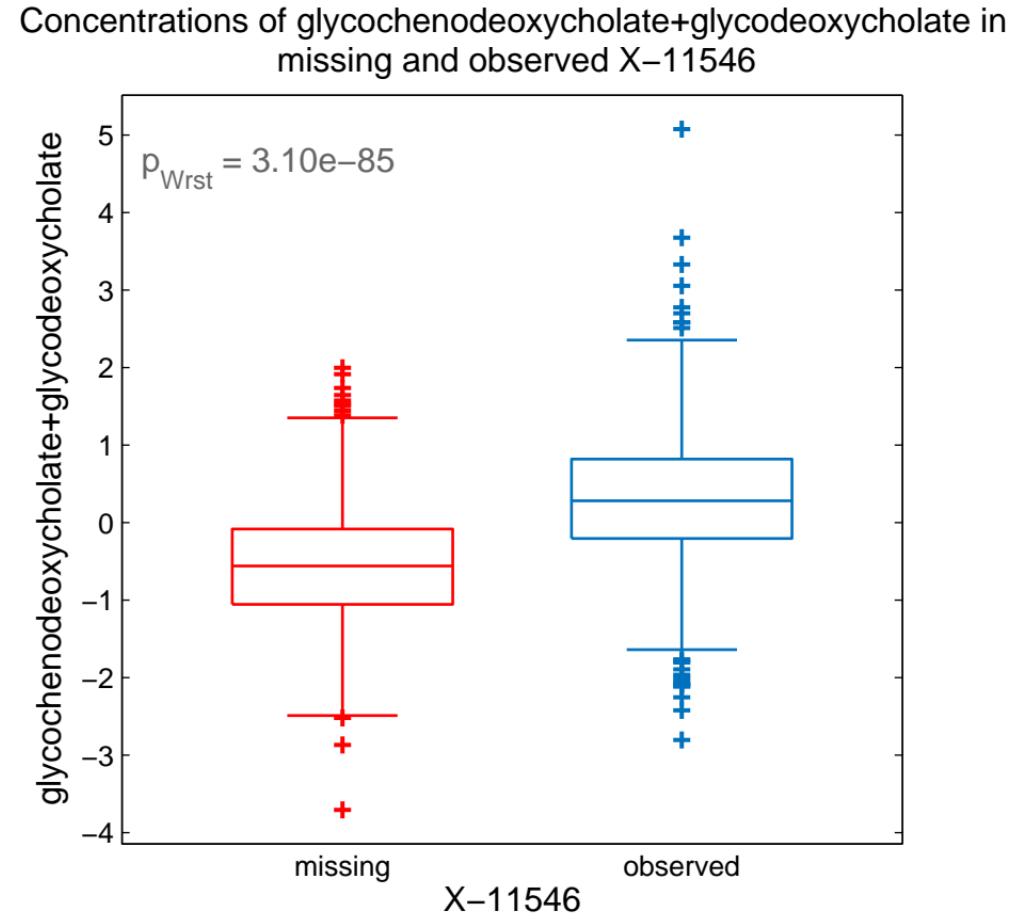
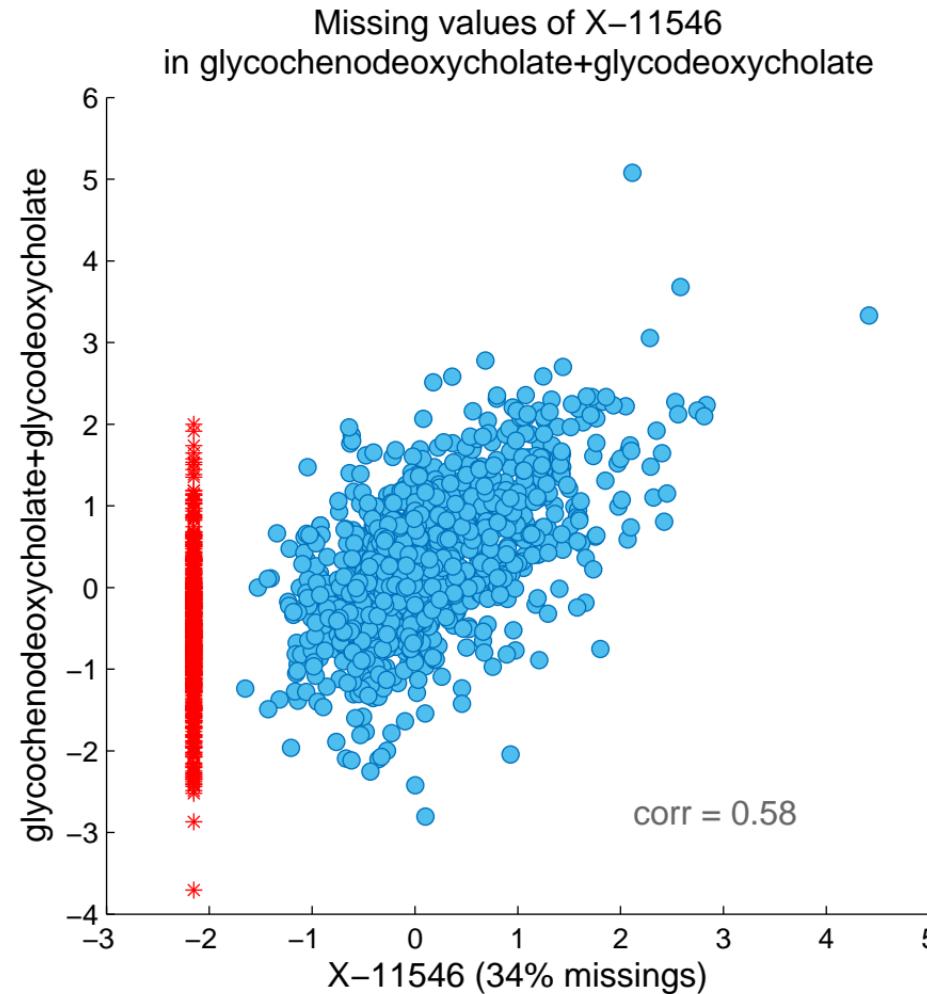


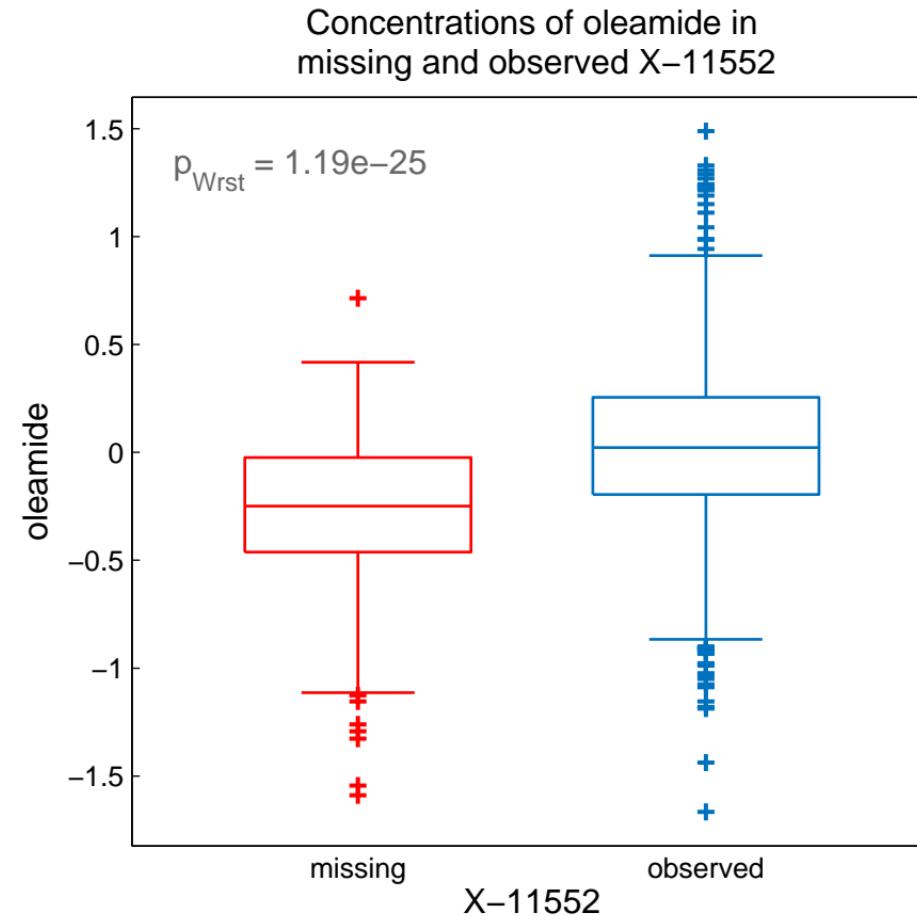
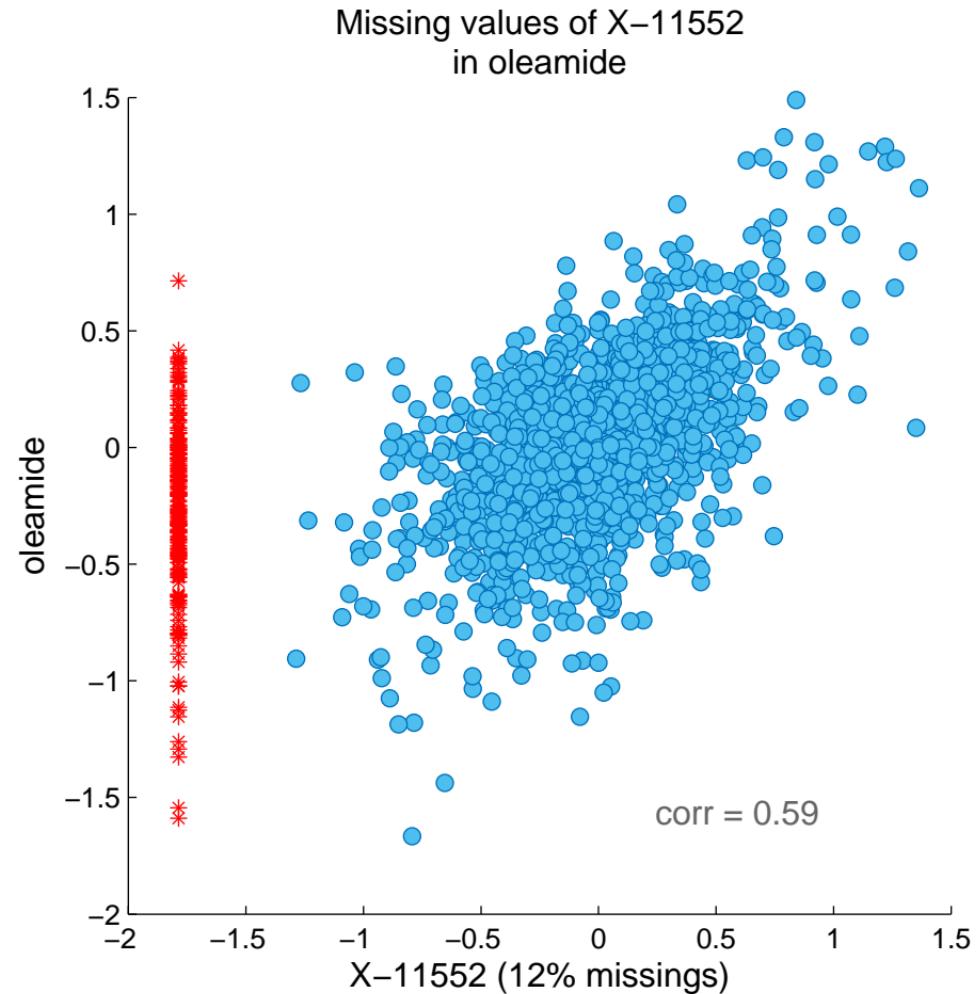


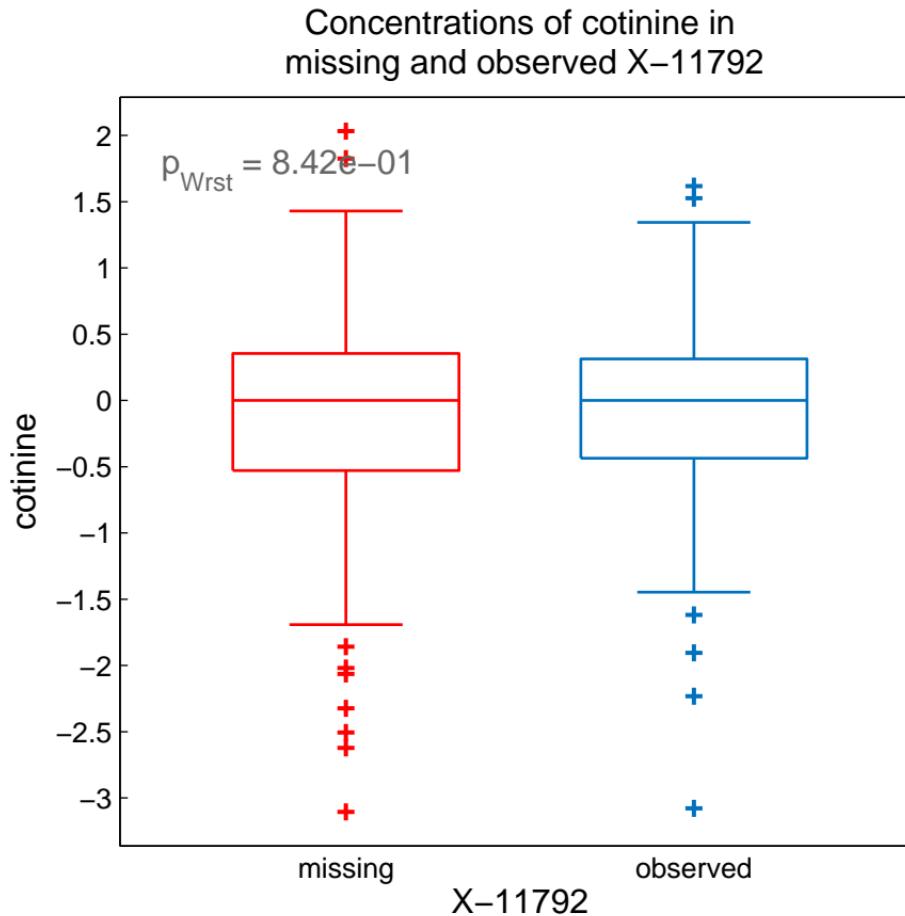
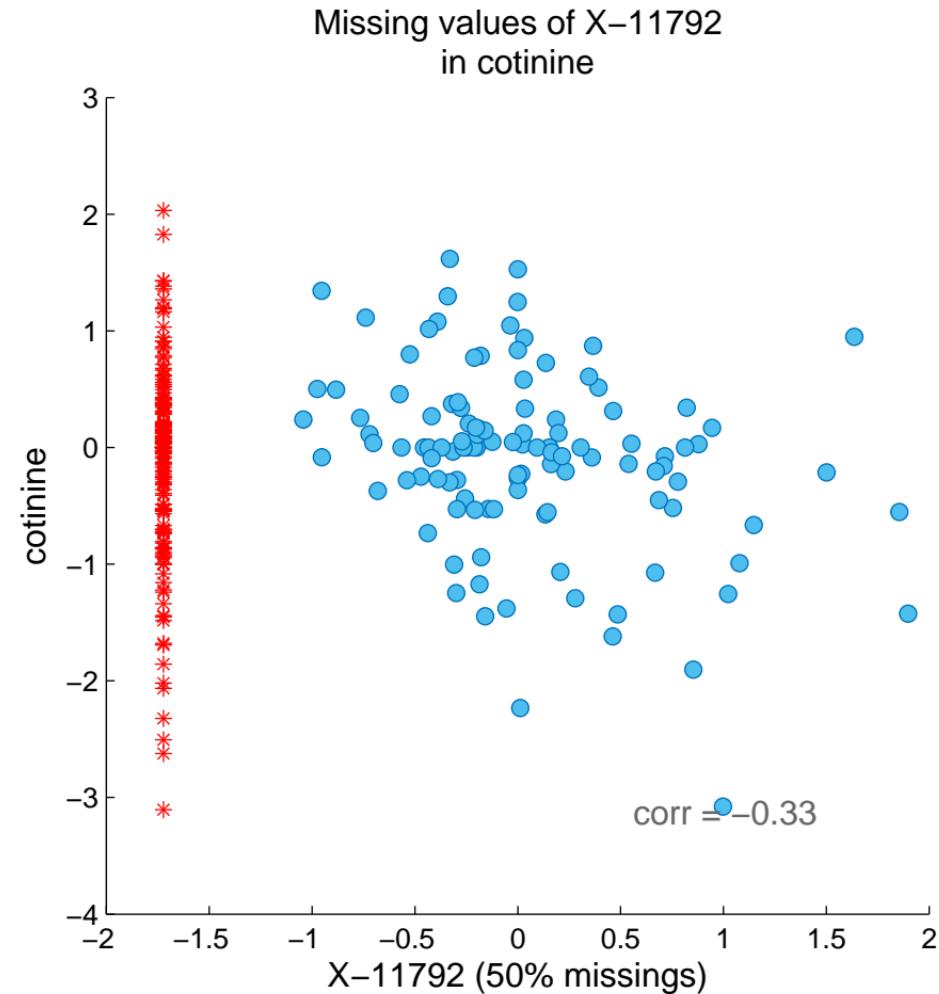


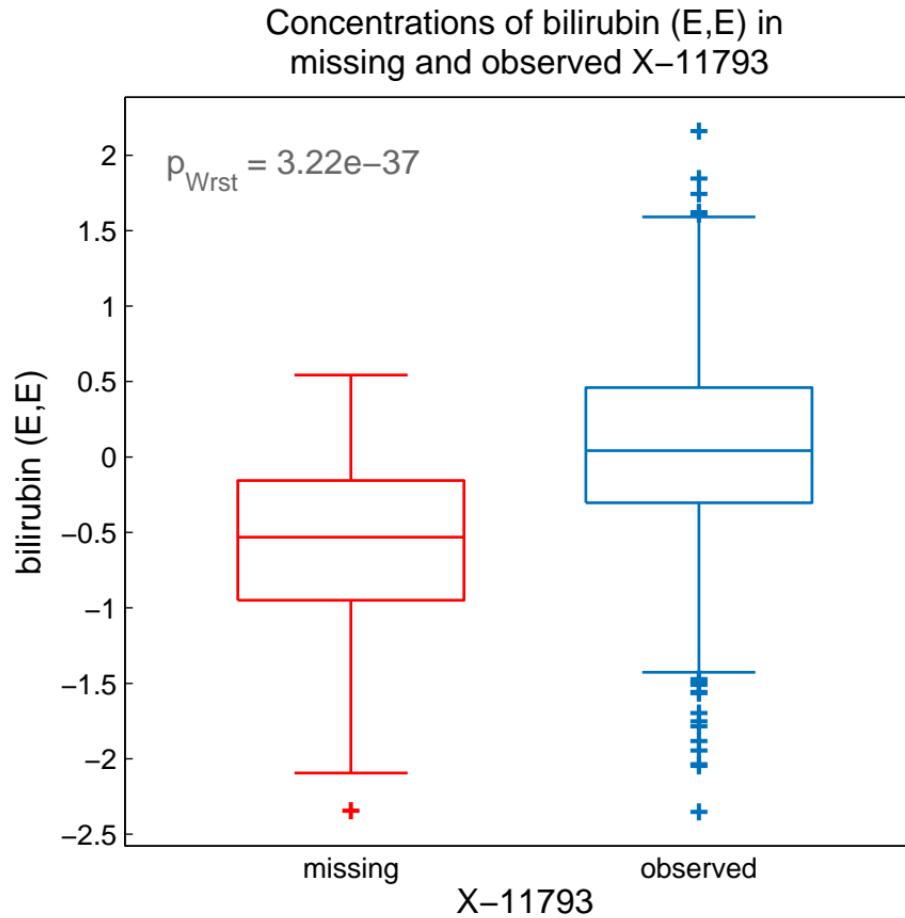
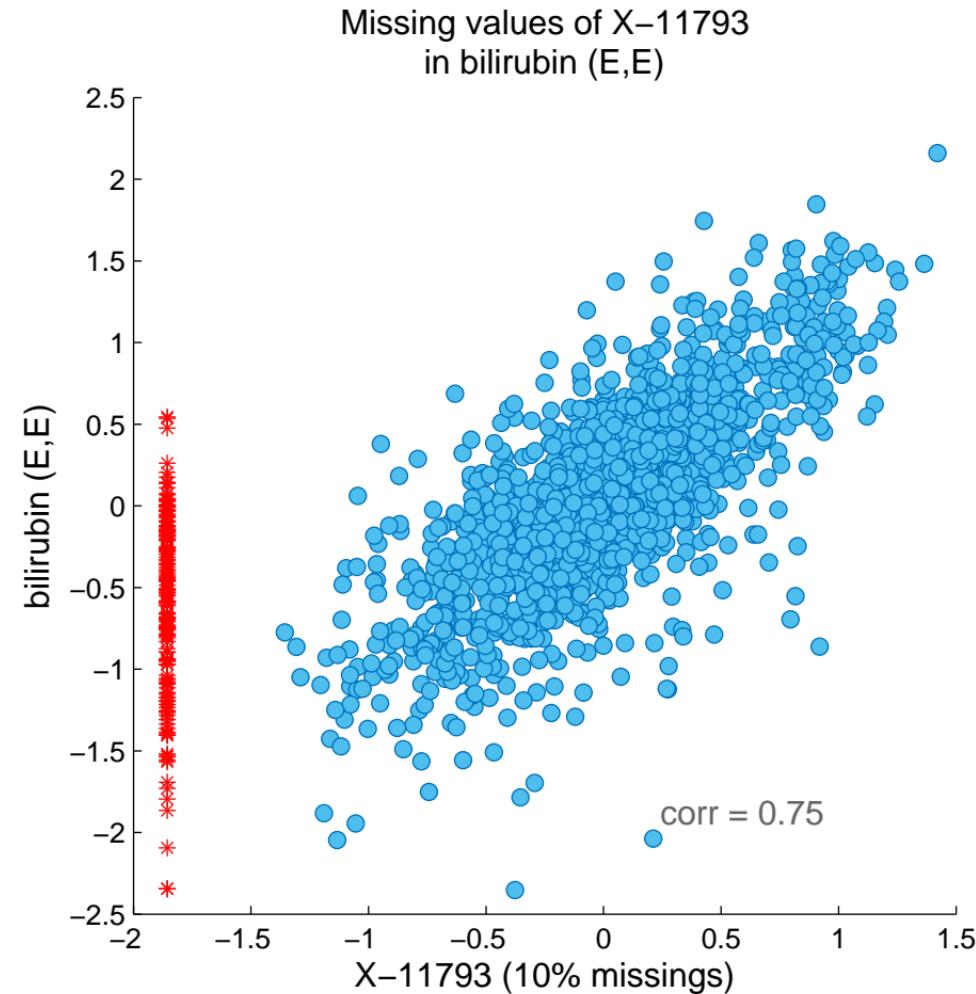


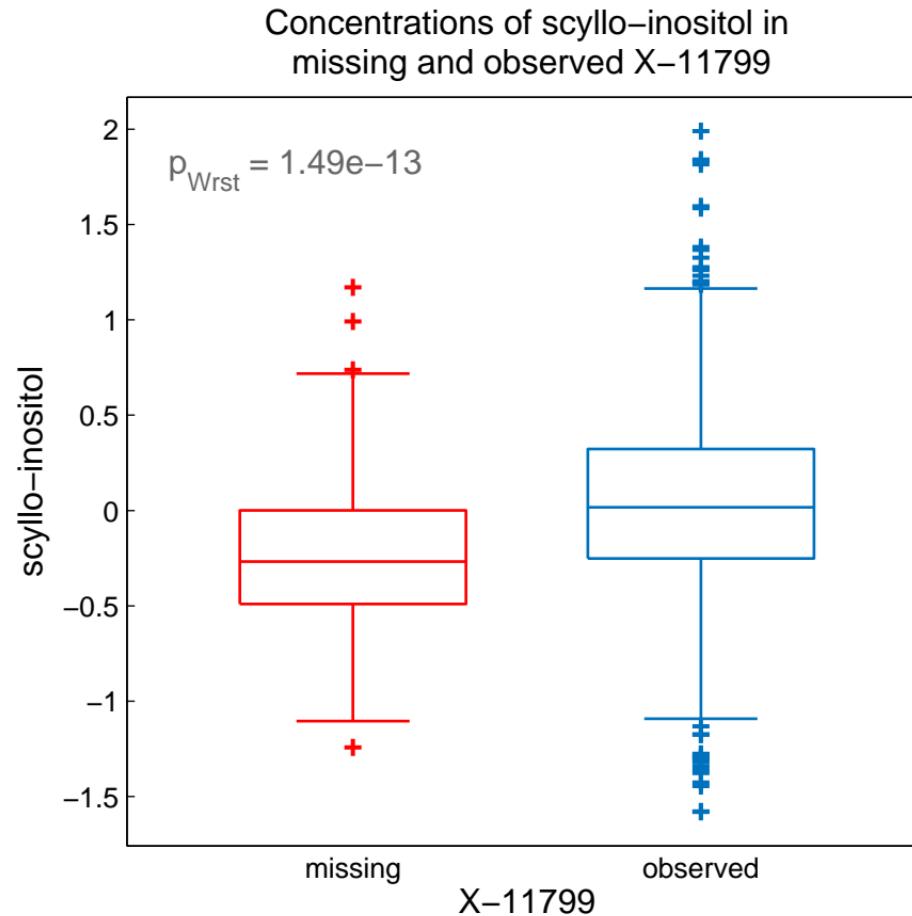
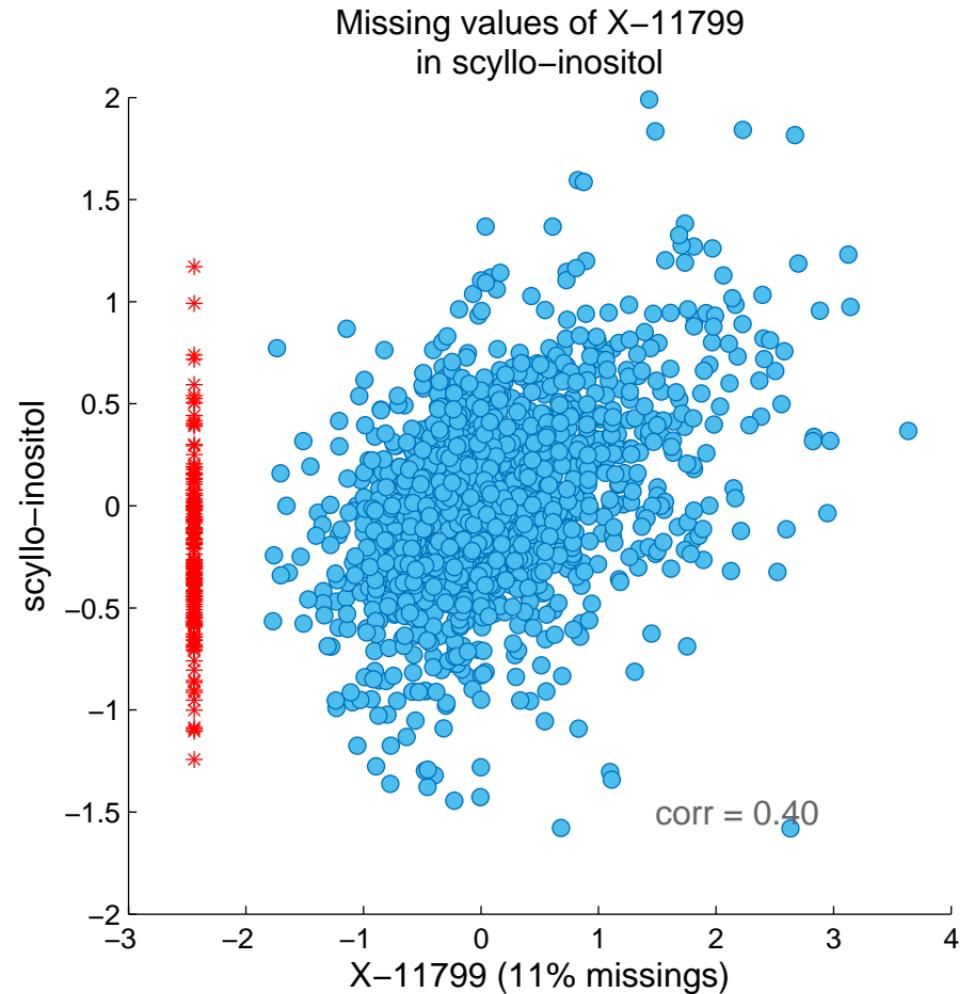


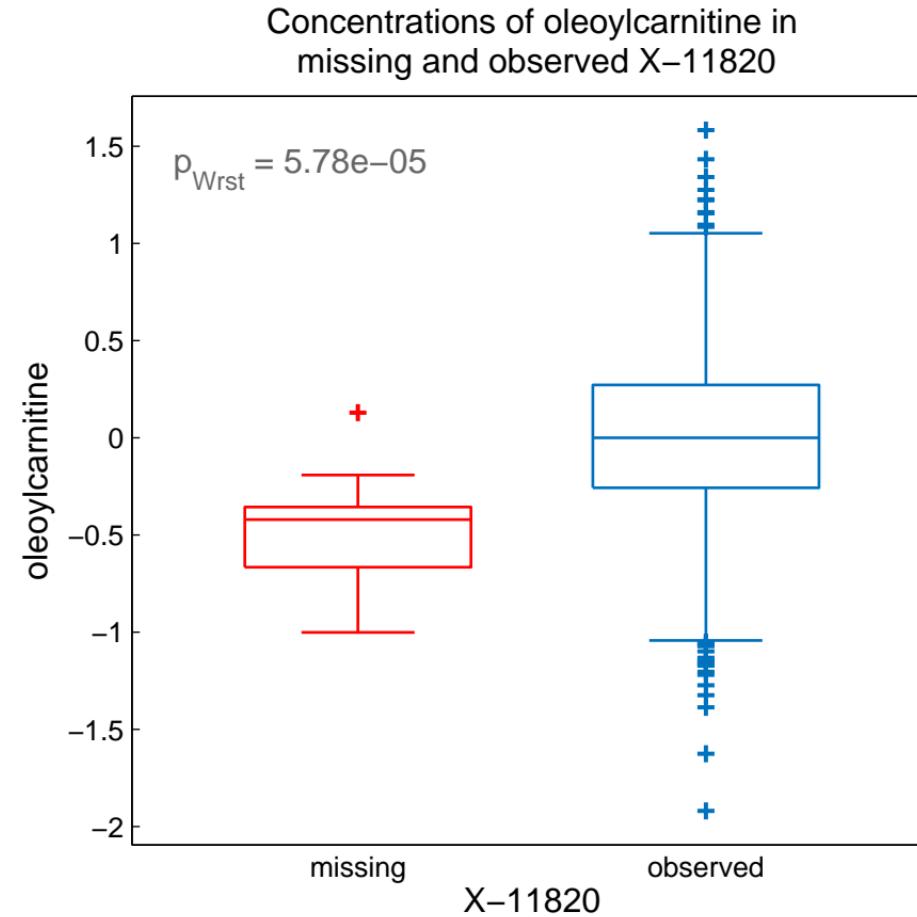
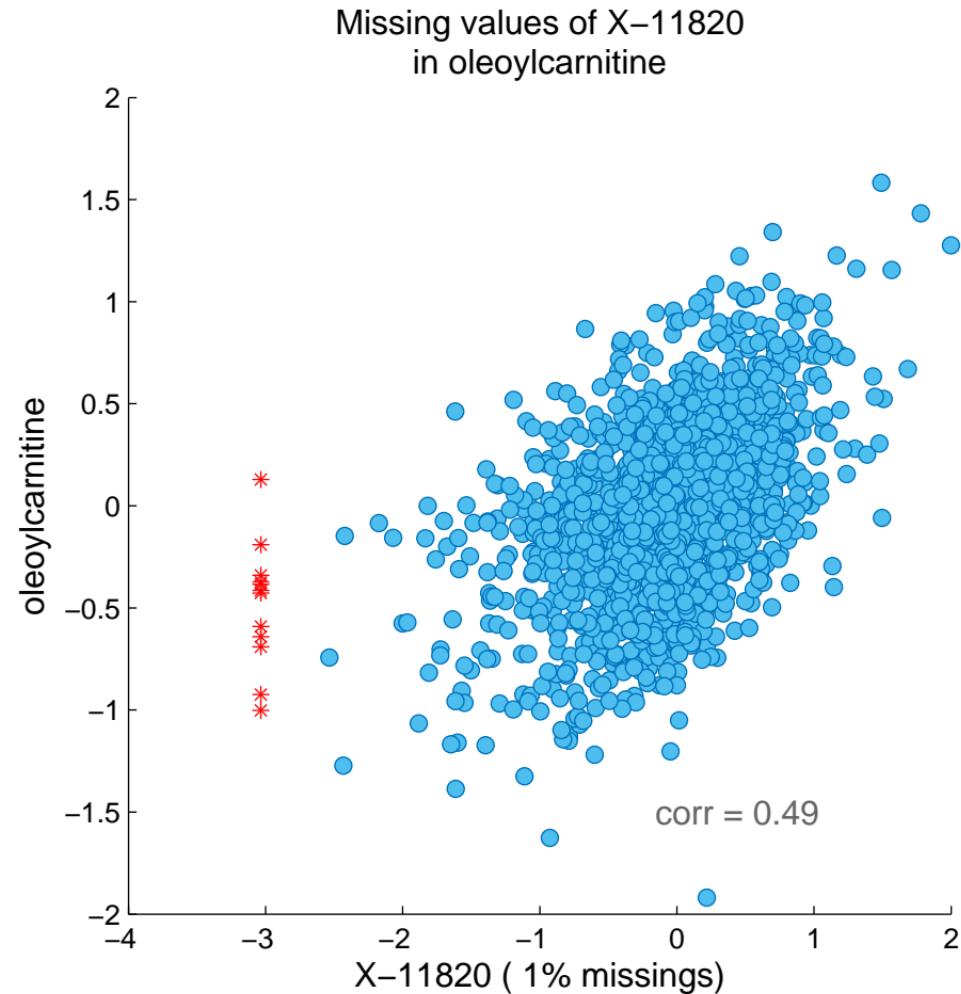


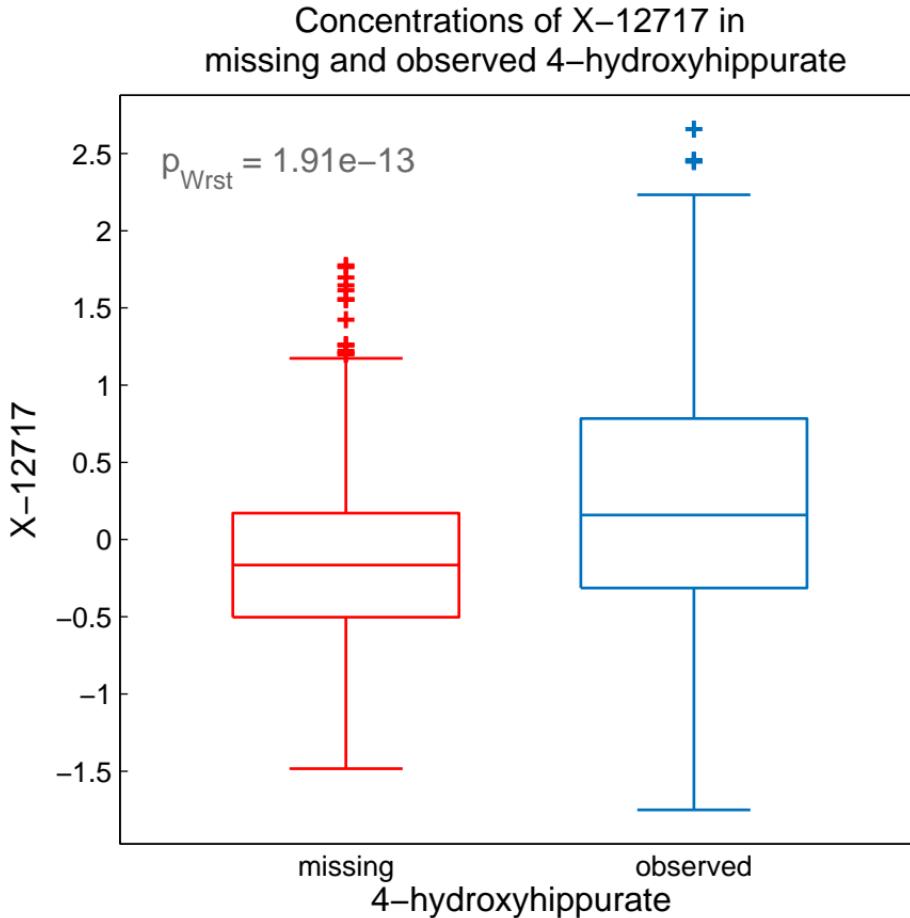
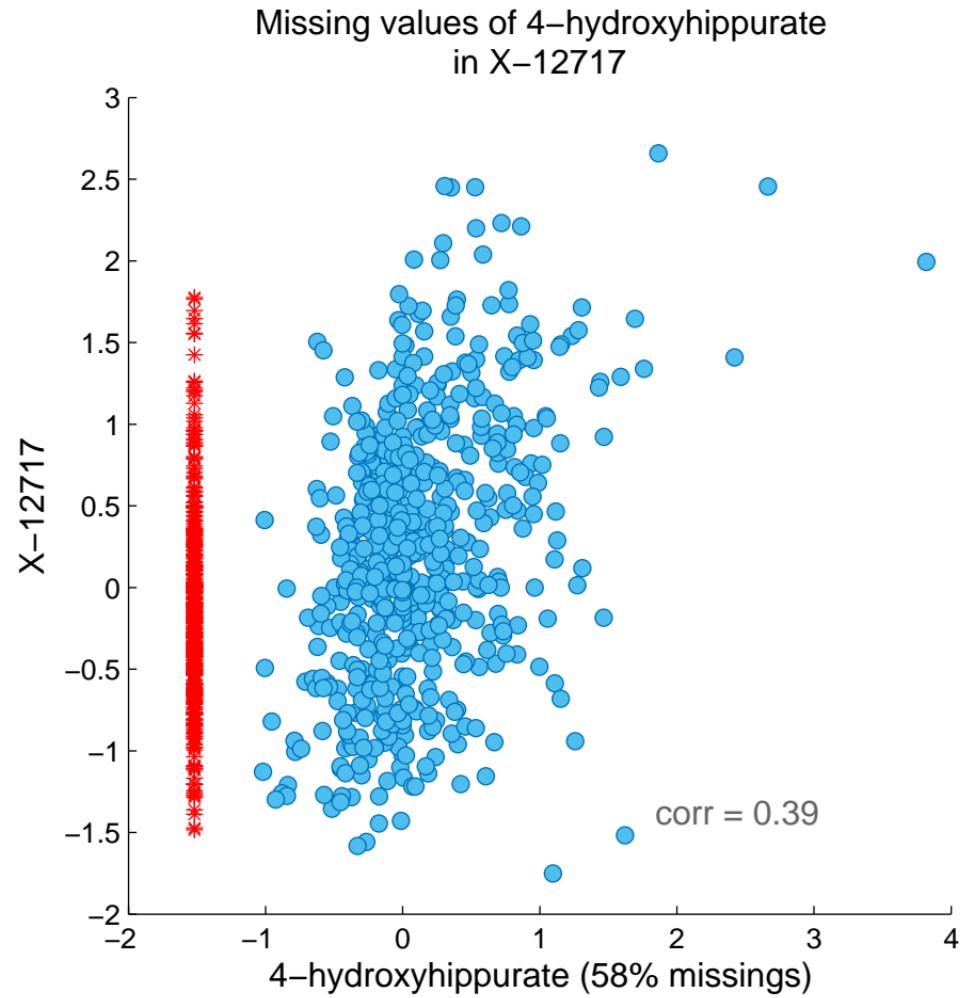




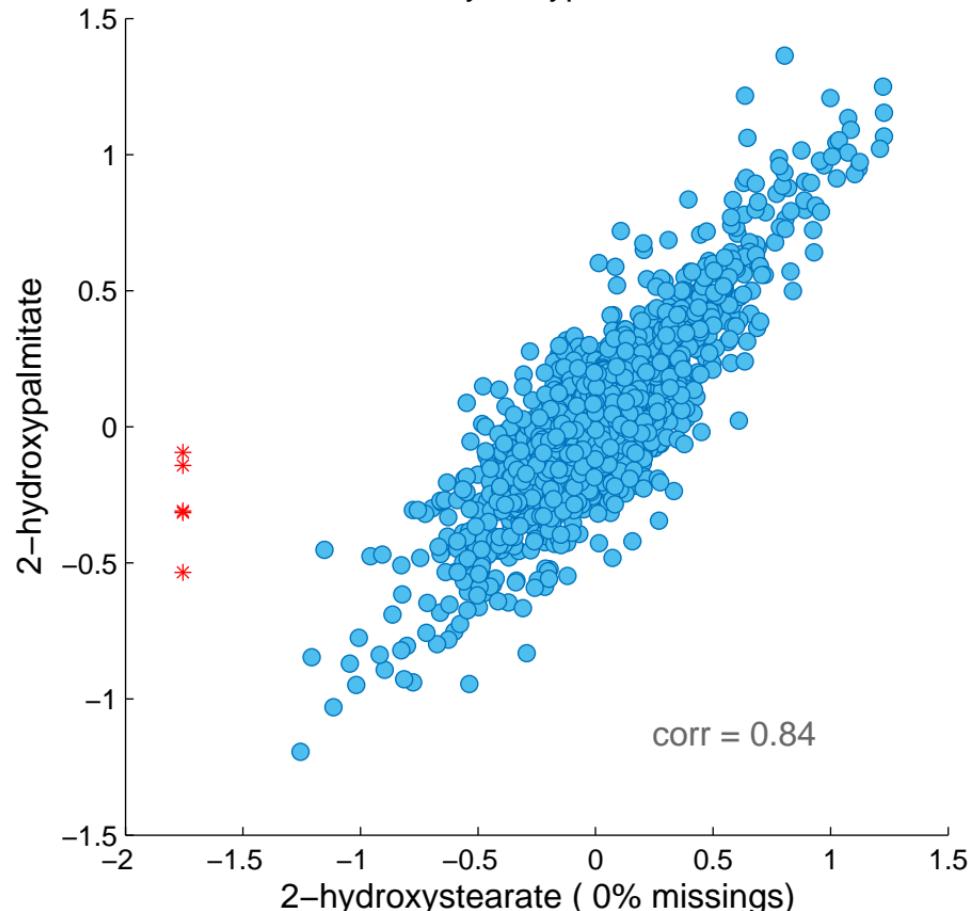




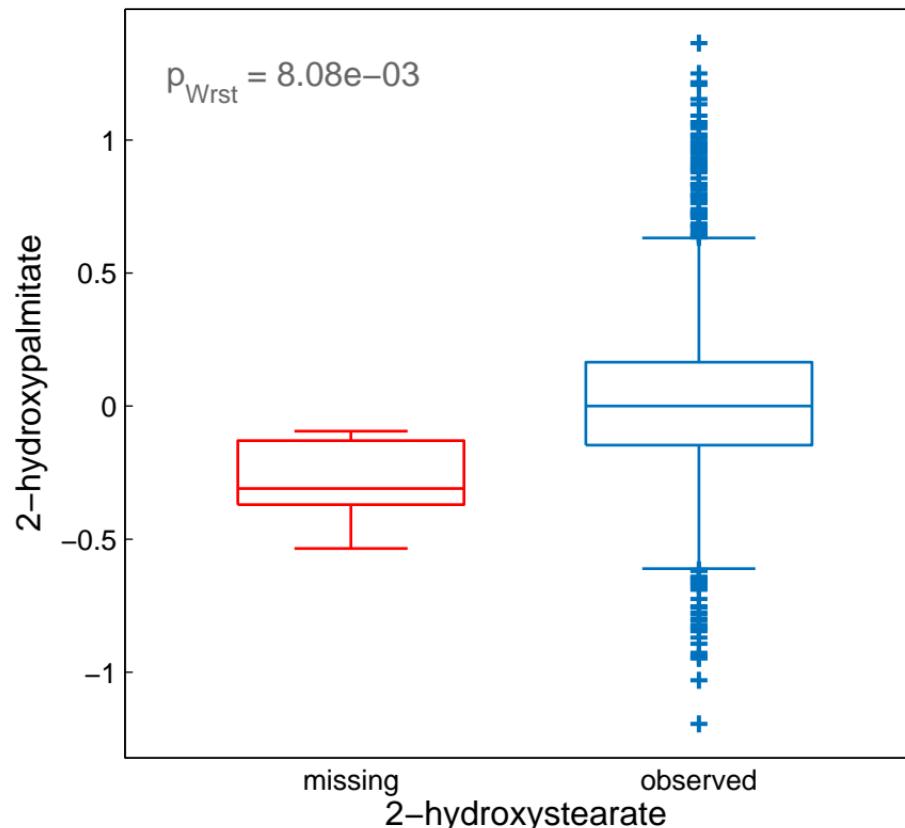


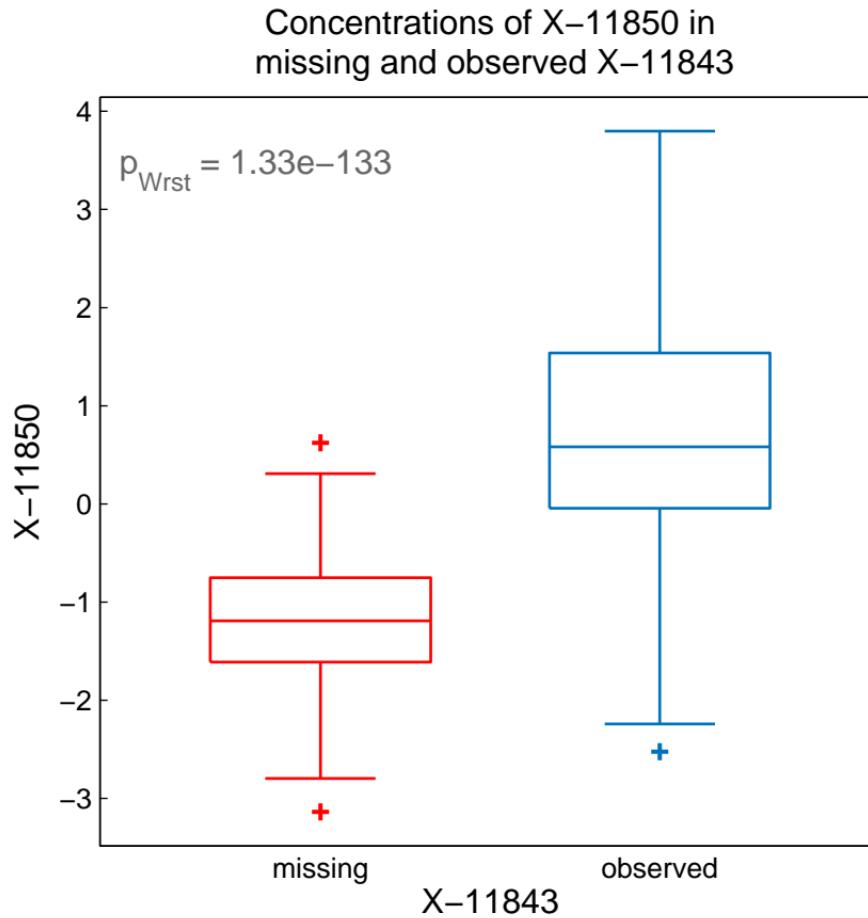
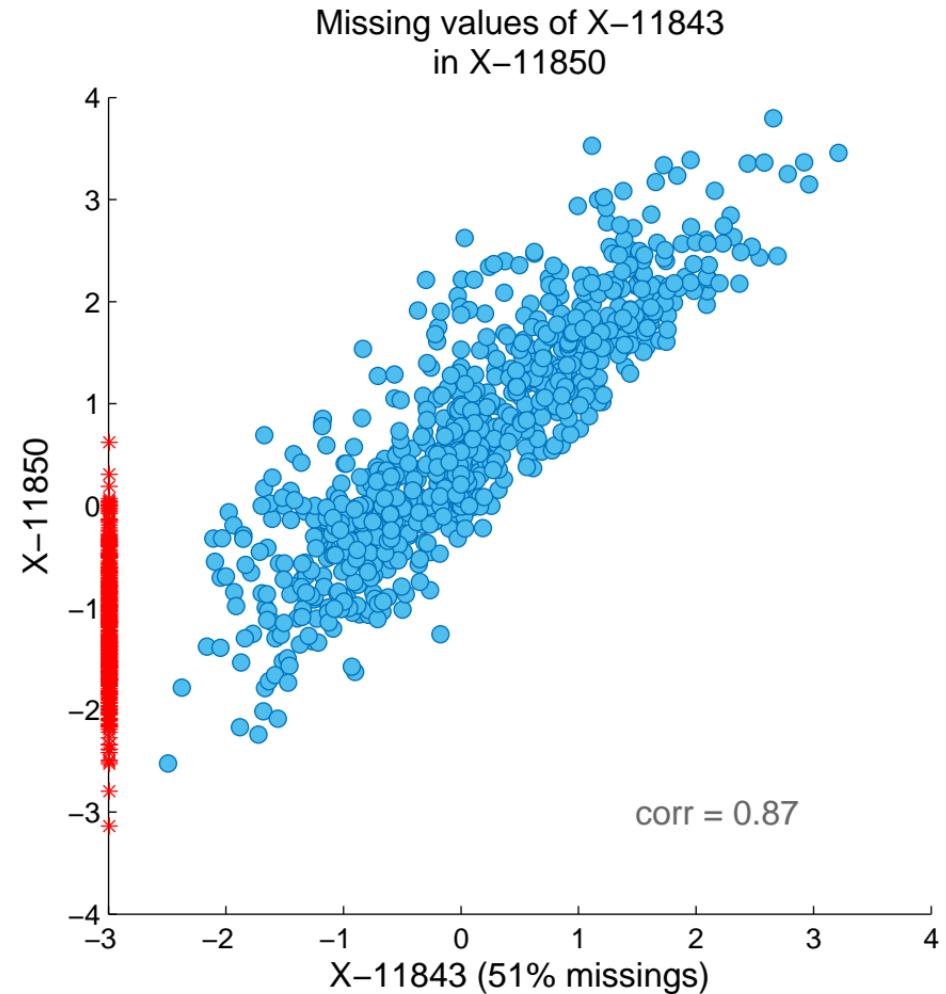


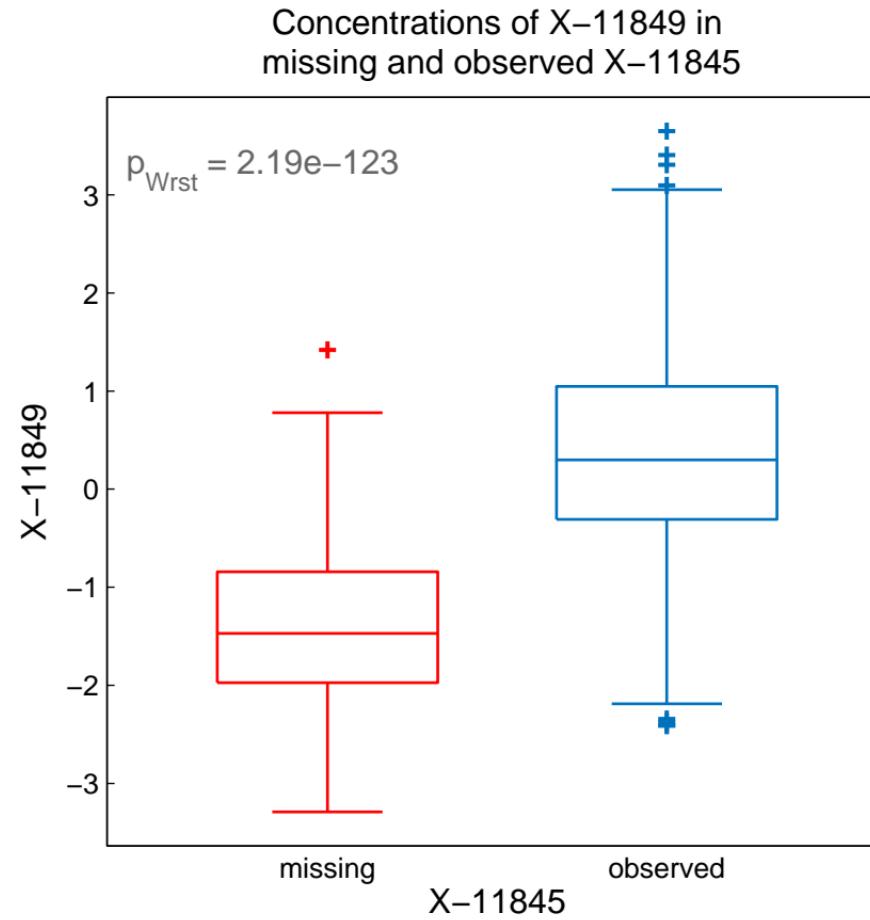
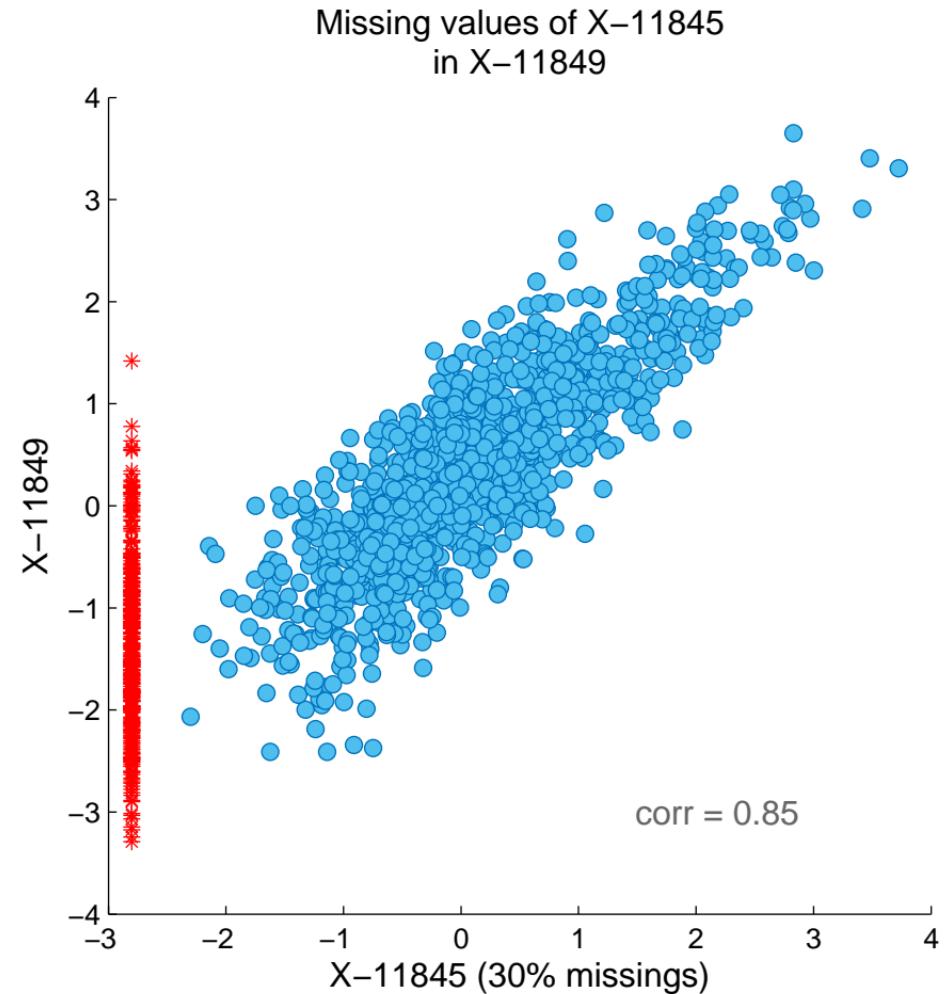
Missing values of 2-hydroxystearate
in 2-hydroxypalmitate

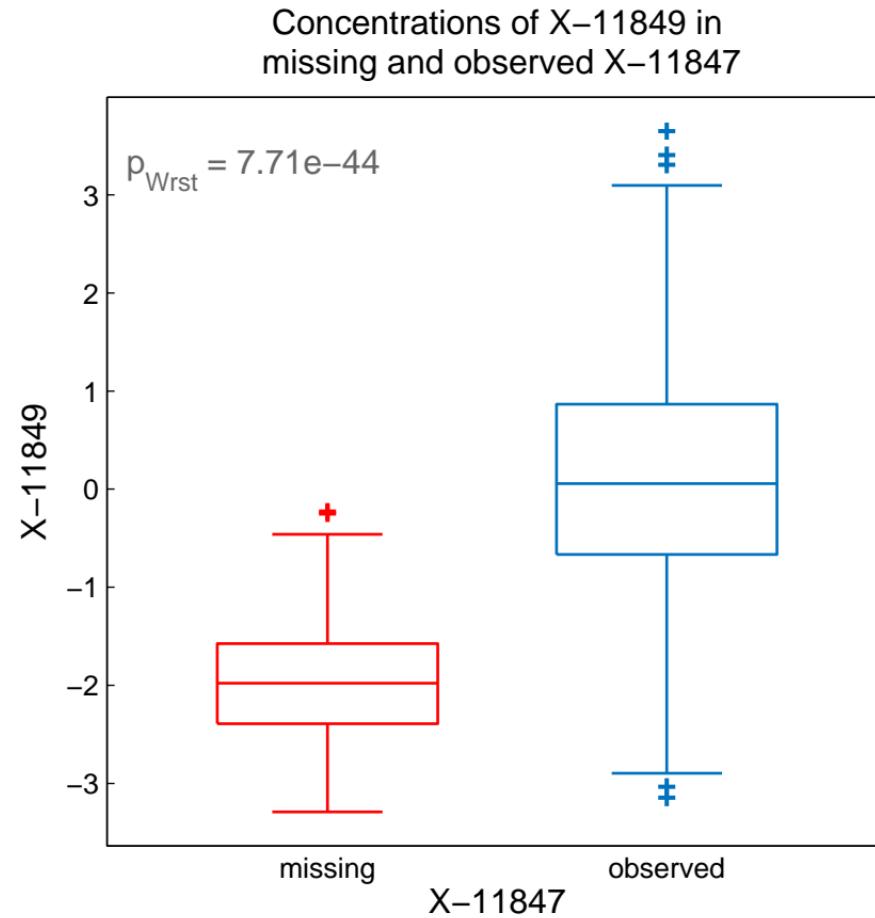
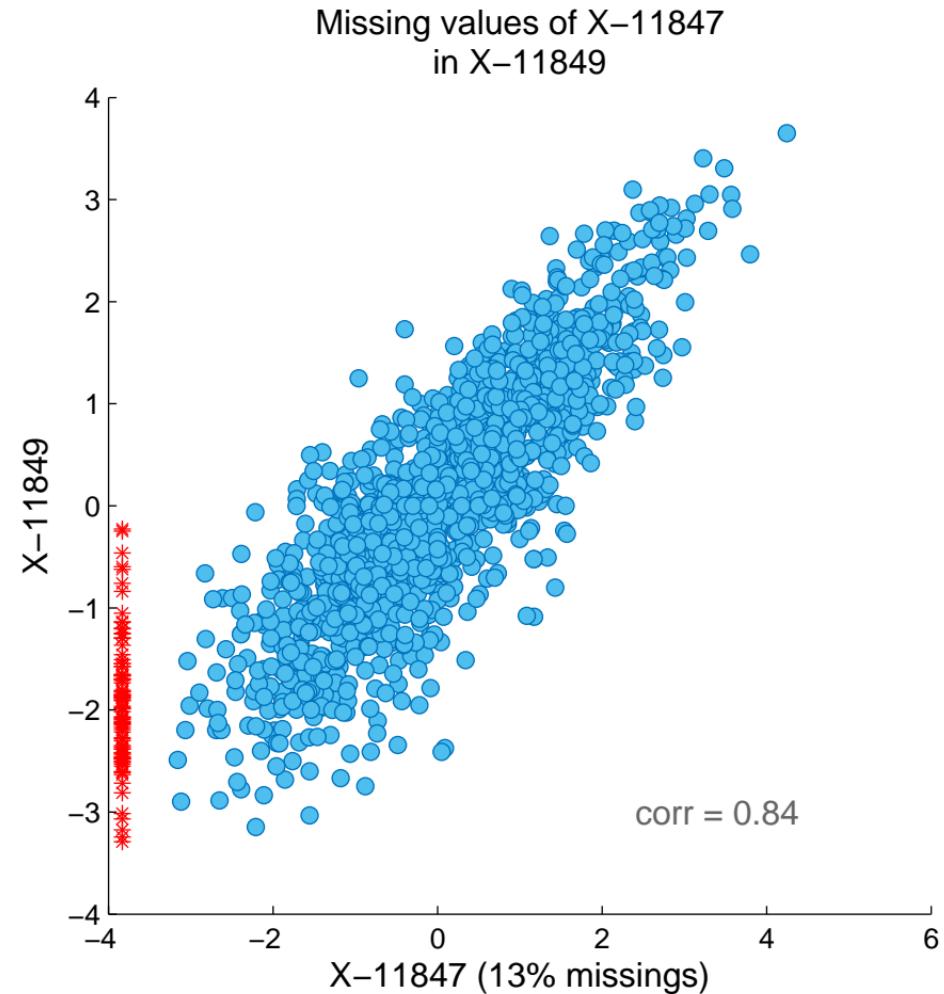


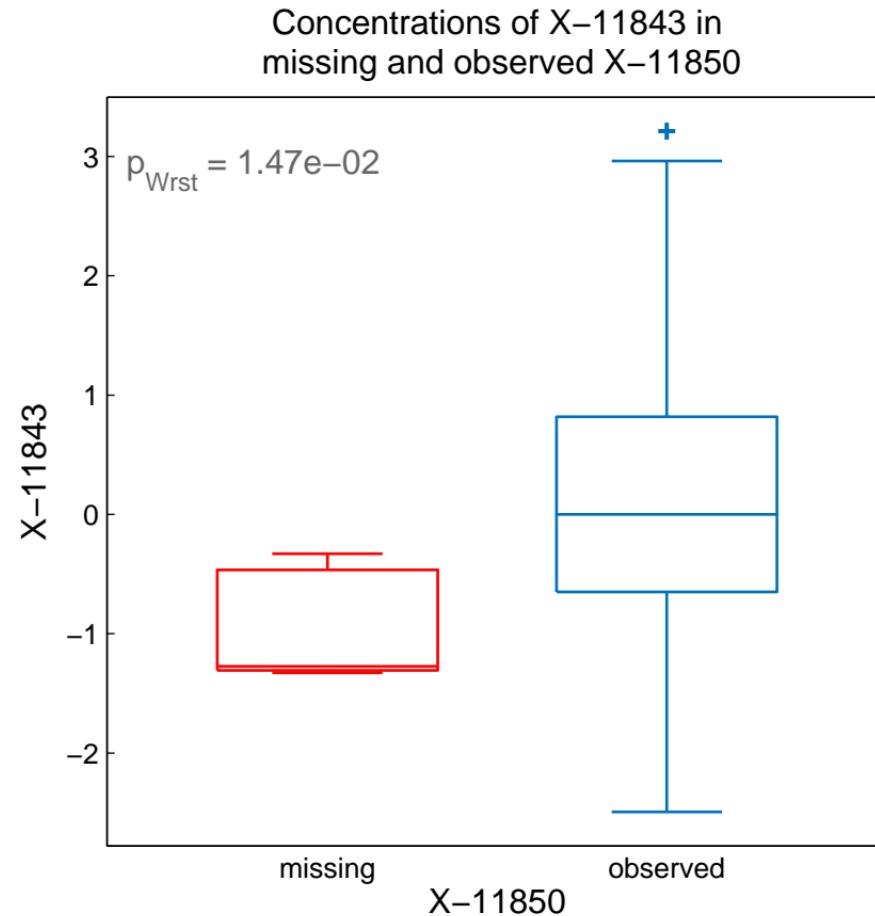
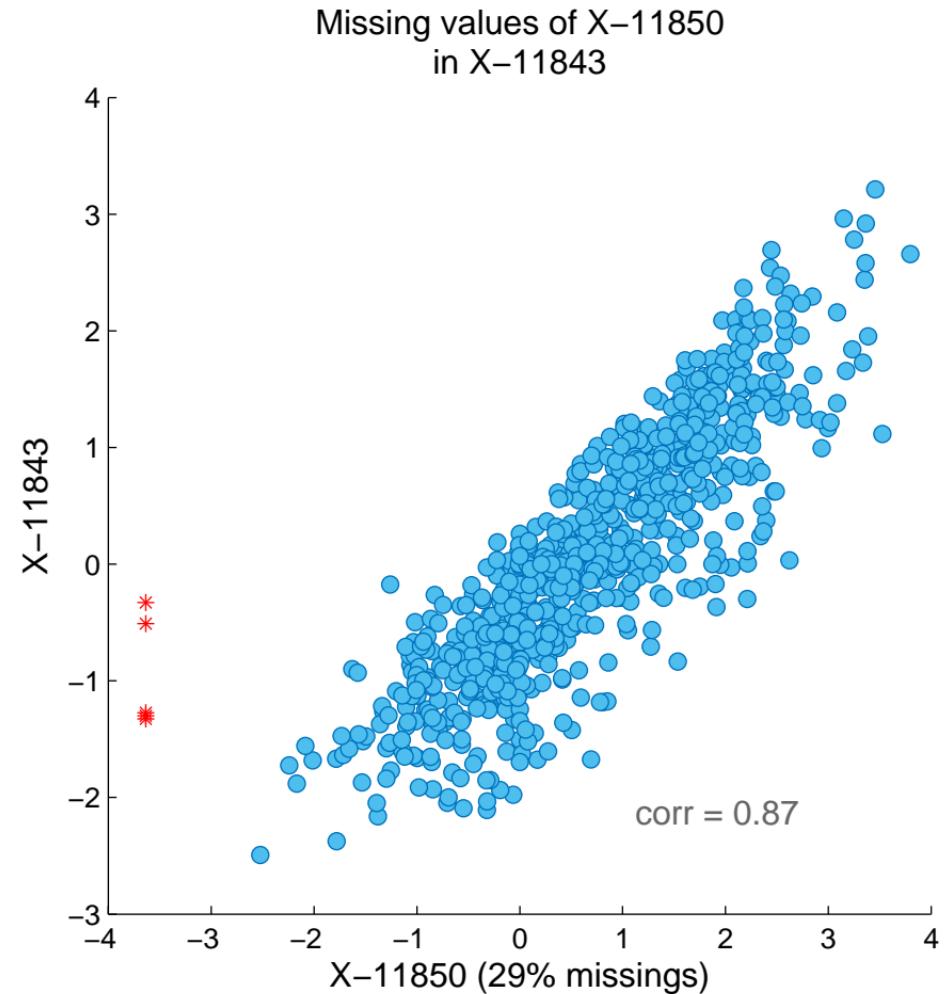
Concentrations of 2-hydroxypalmitate in
missing and observed 2-hydroxystearate

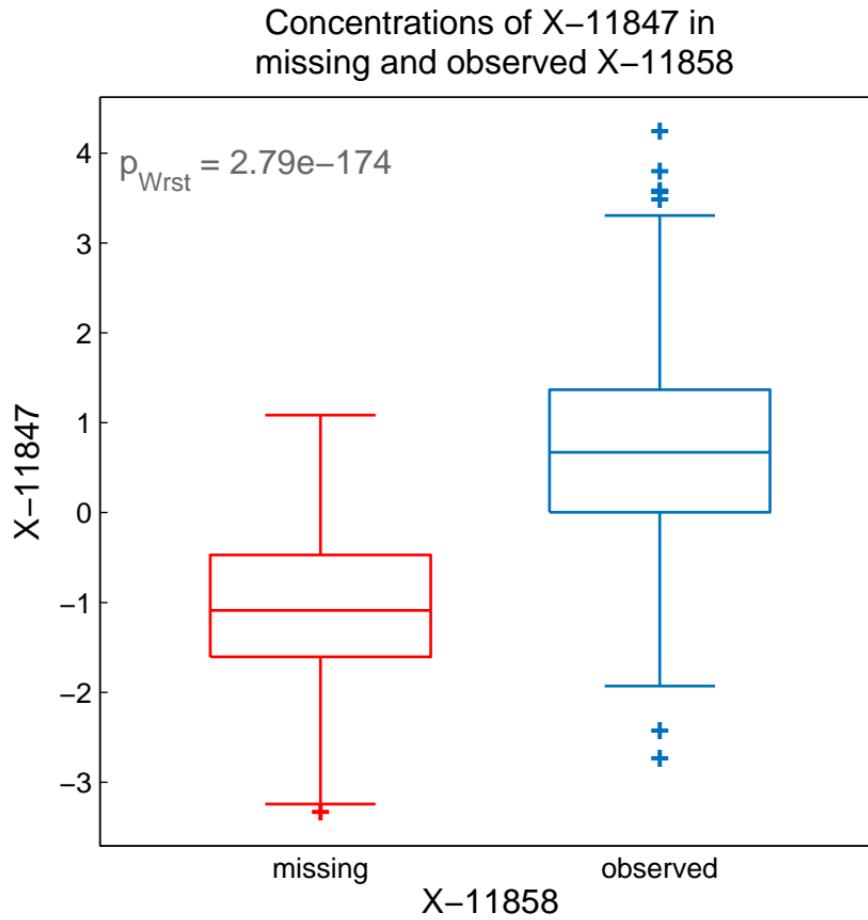
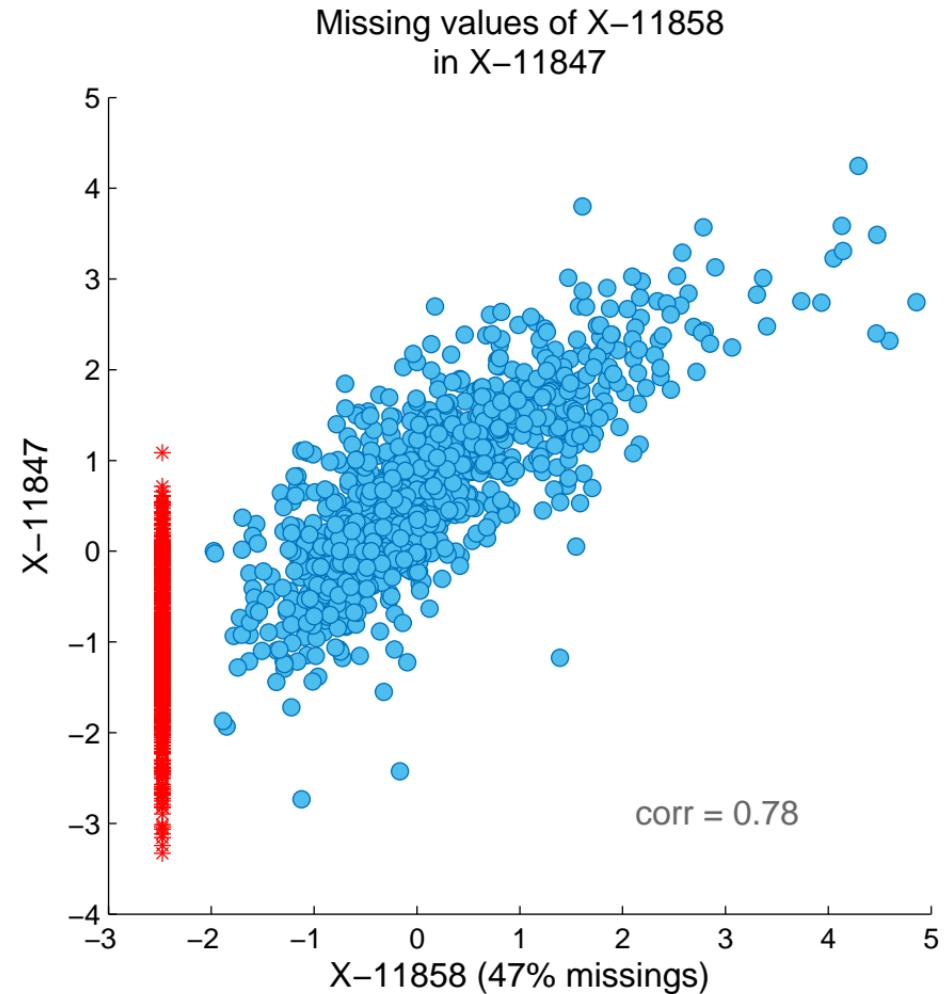


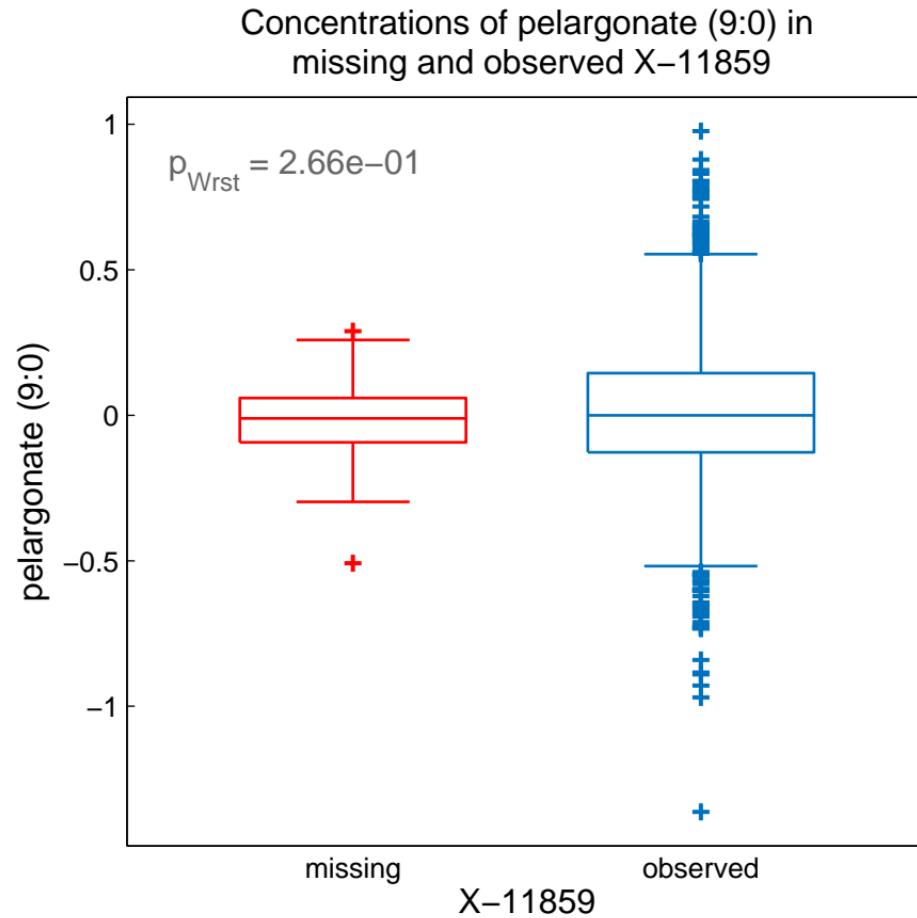
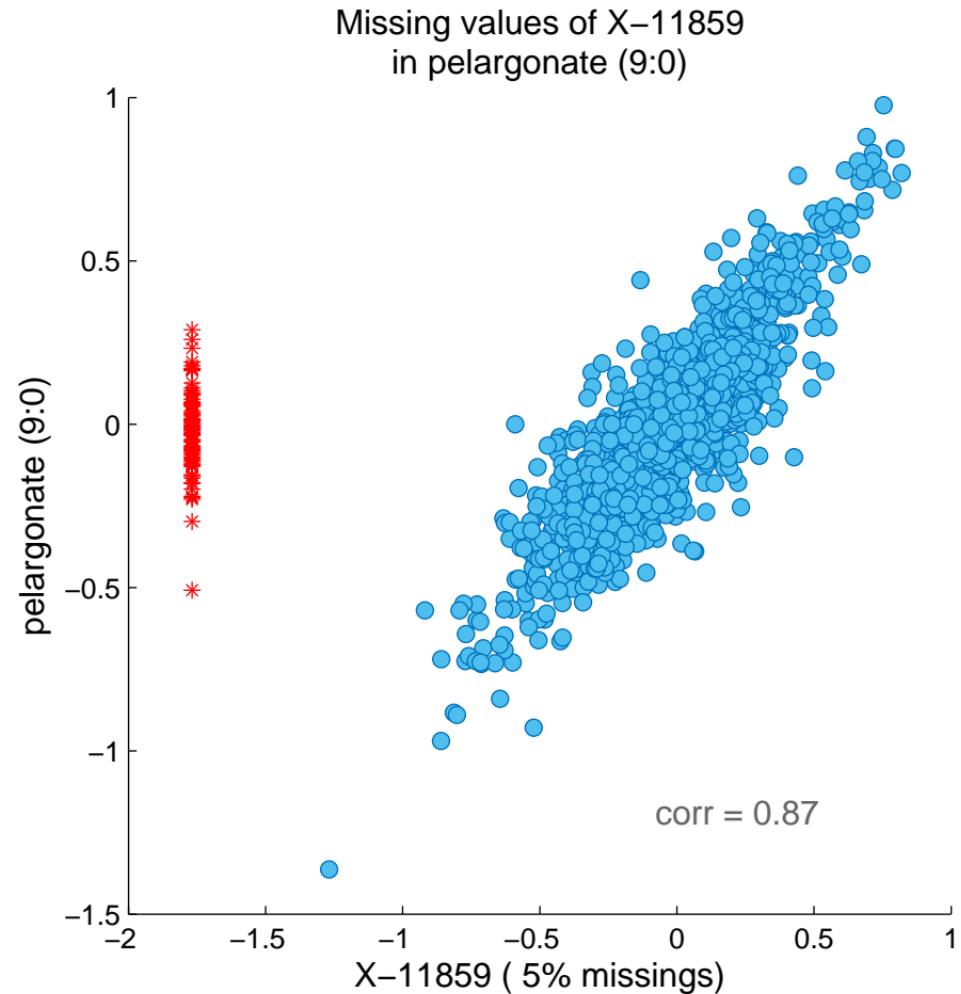


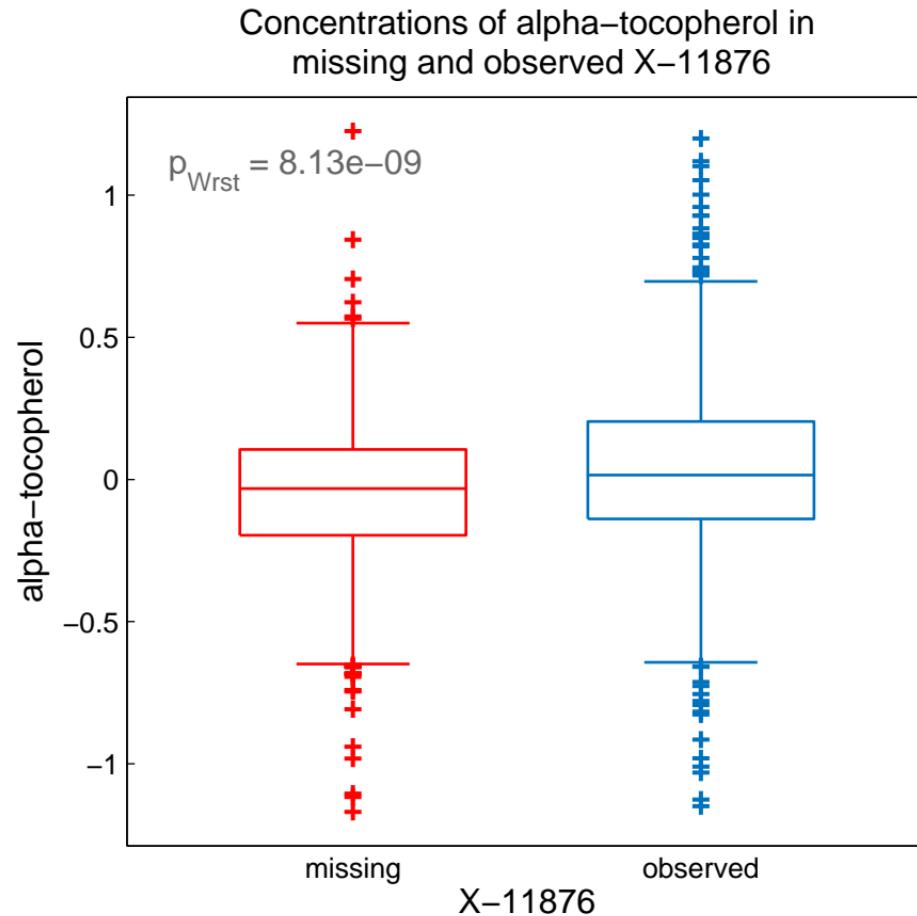
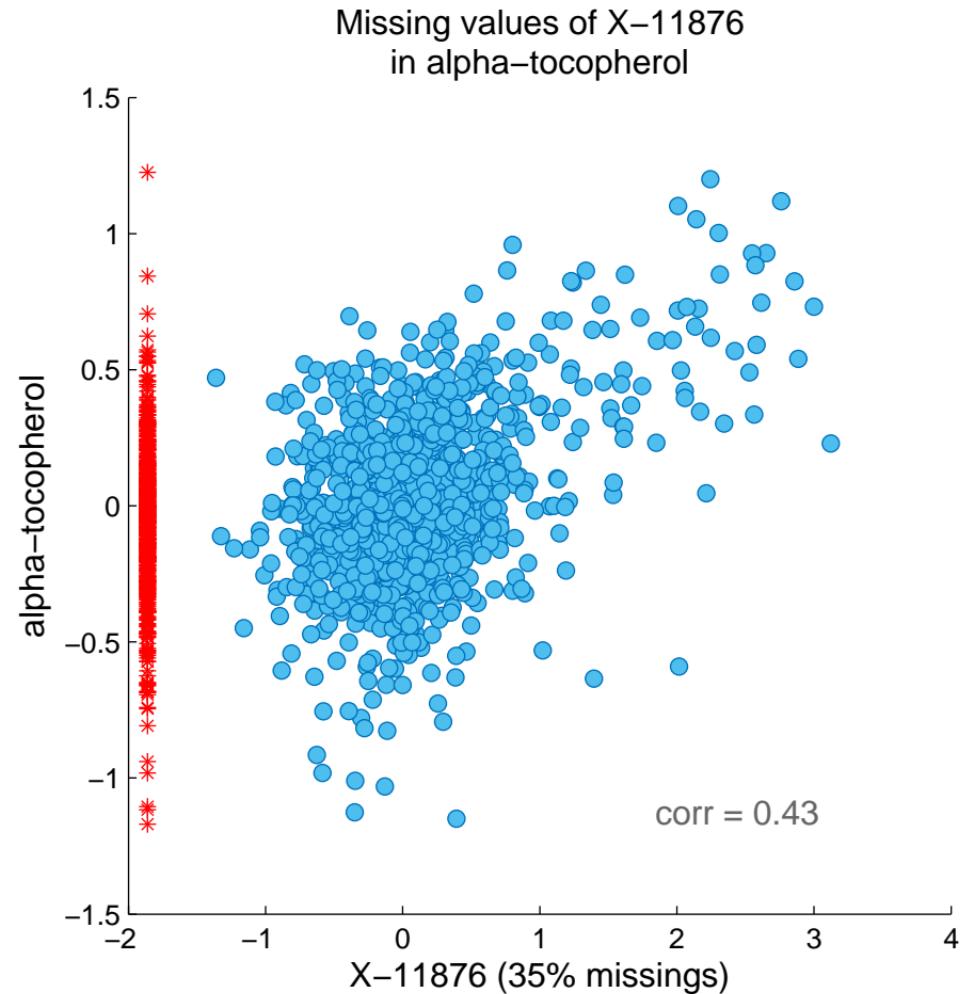


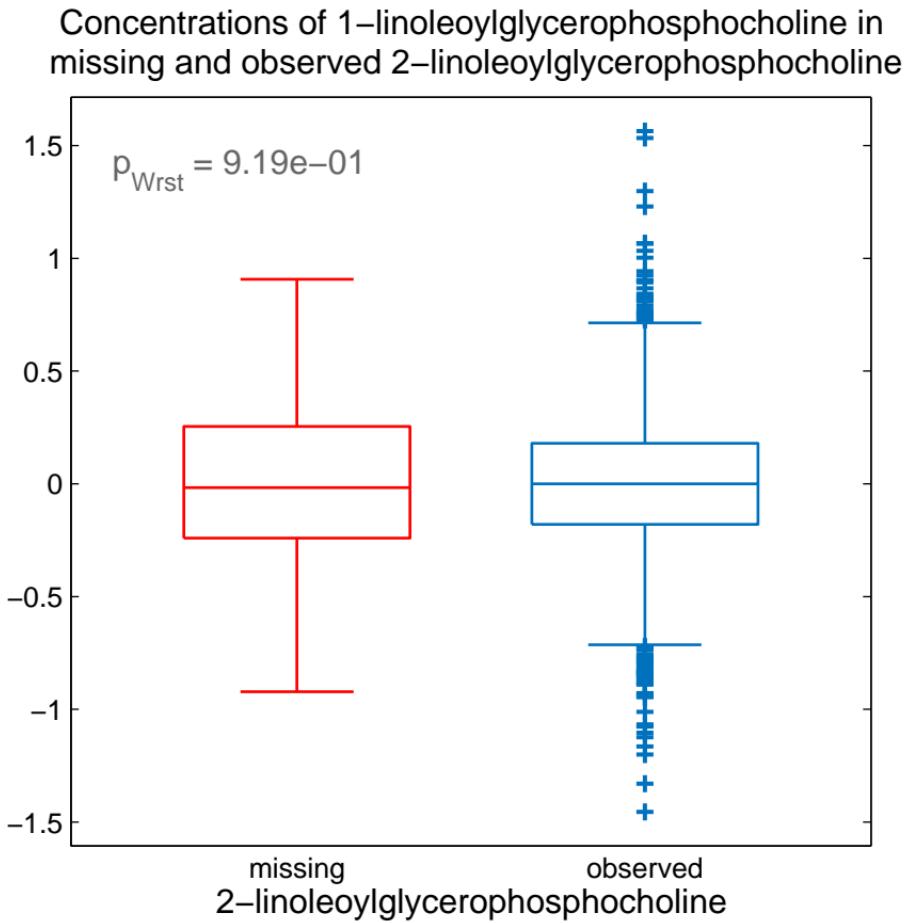
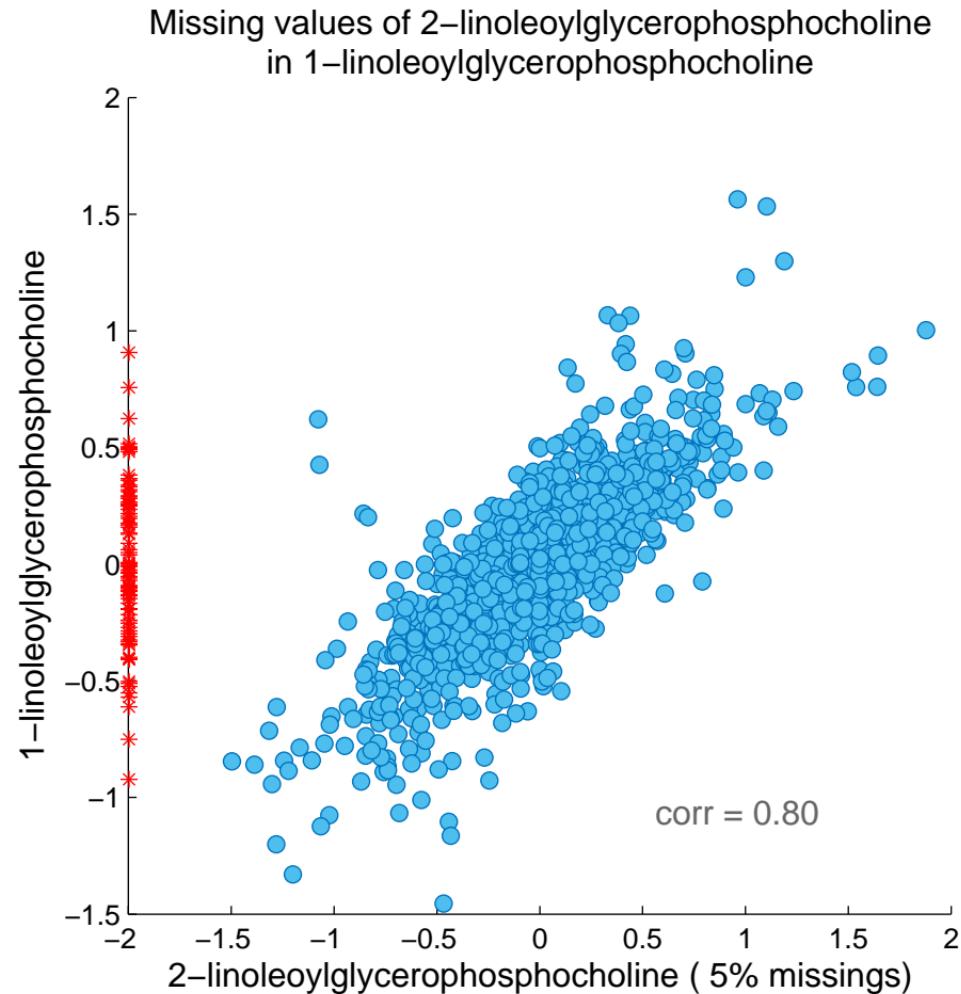


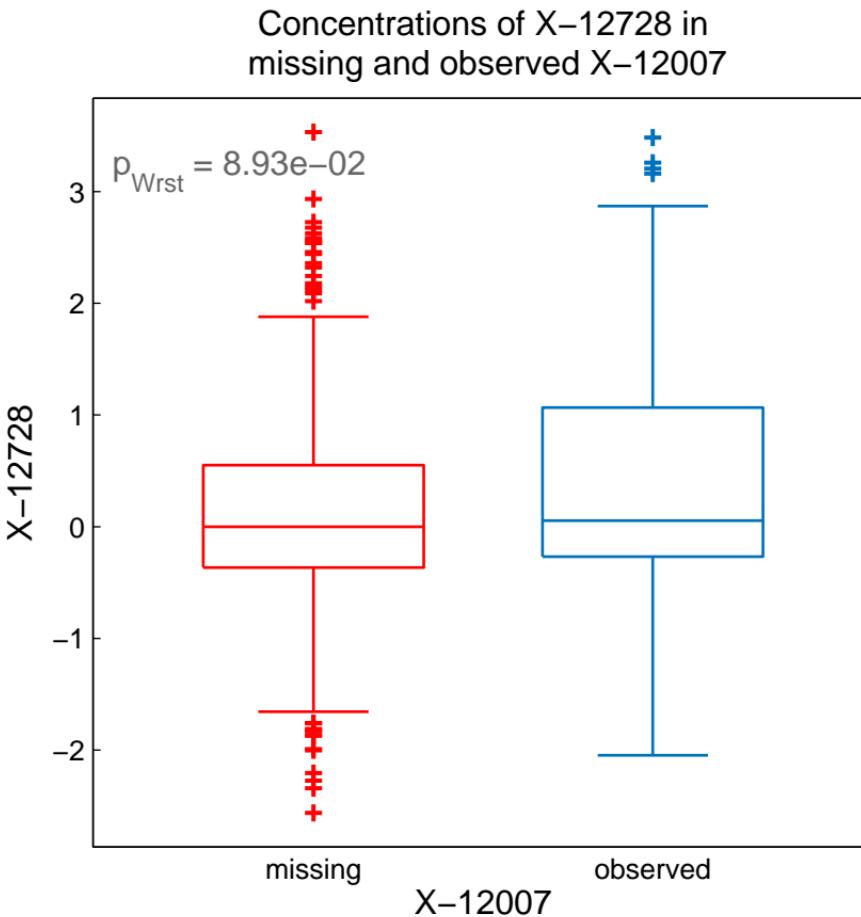
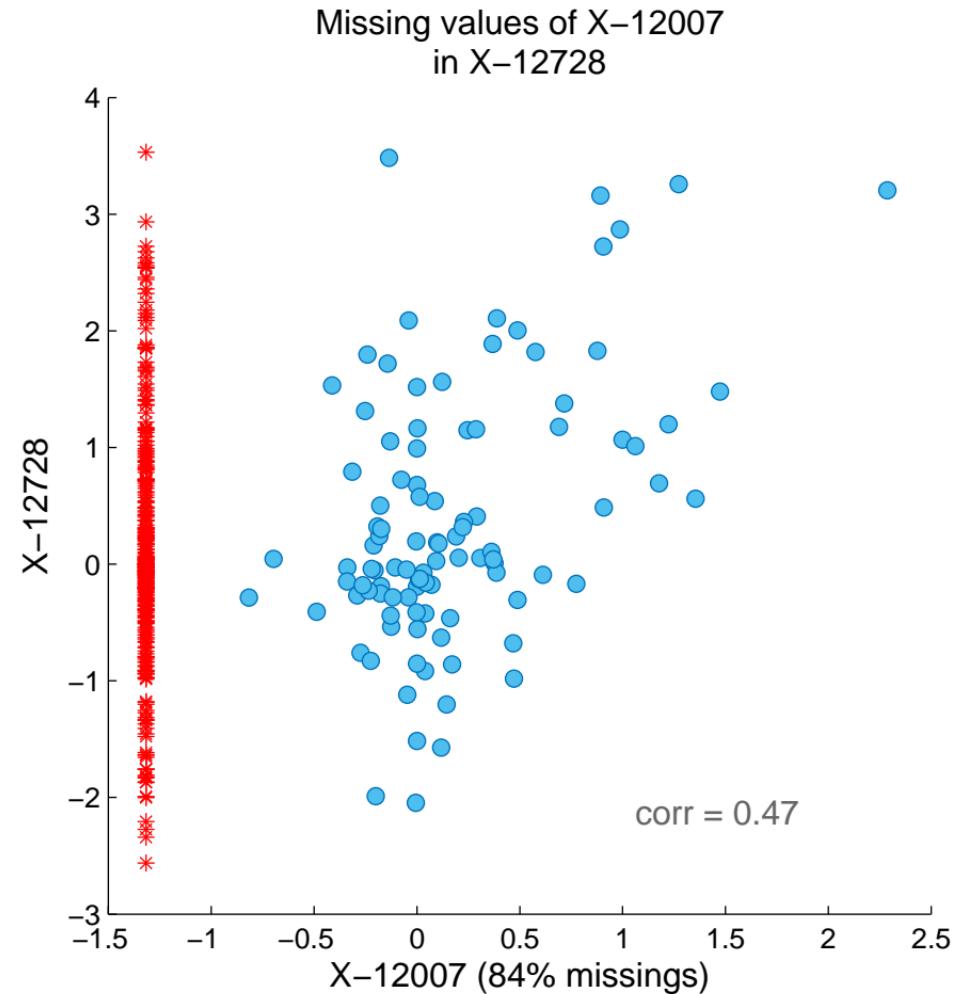


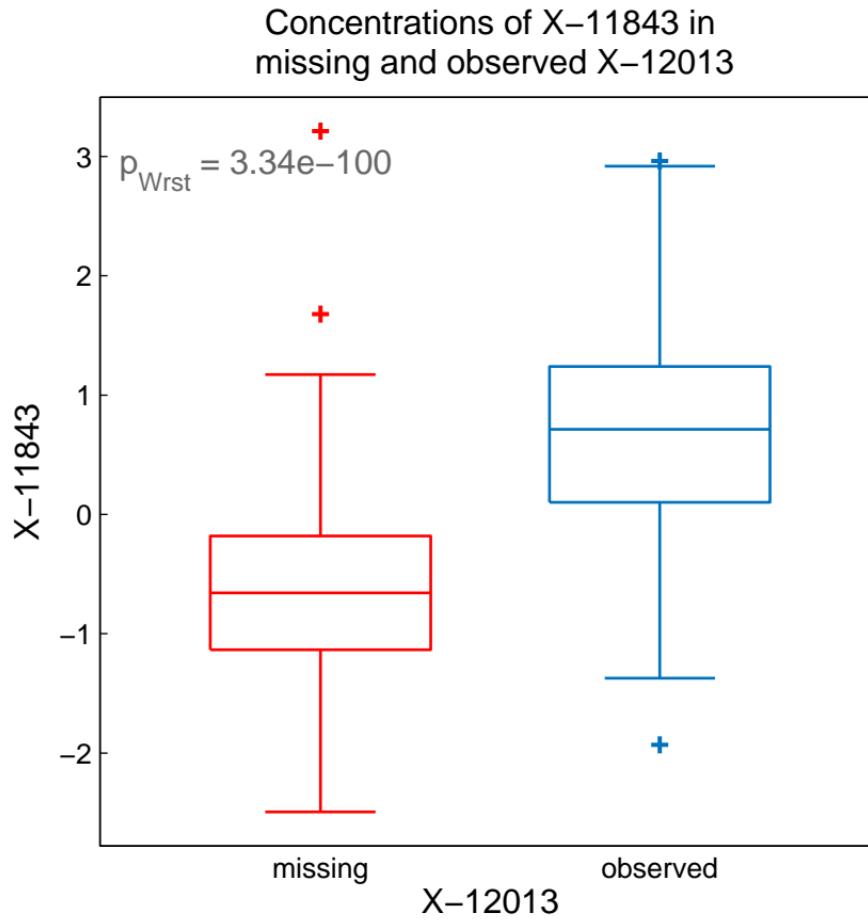
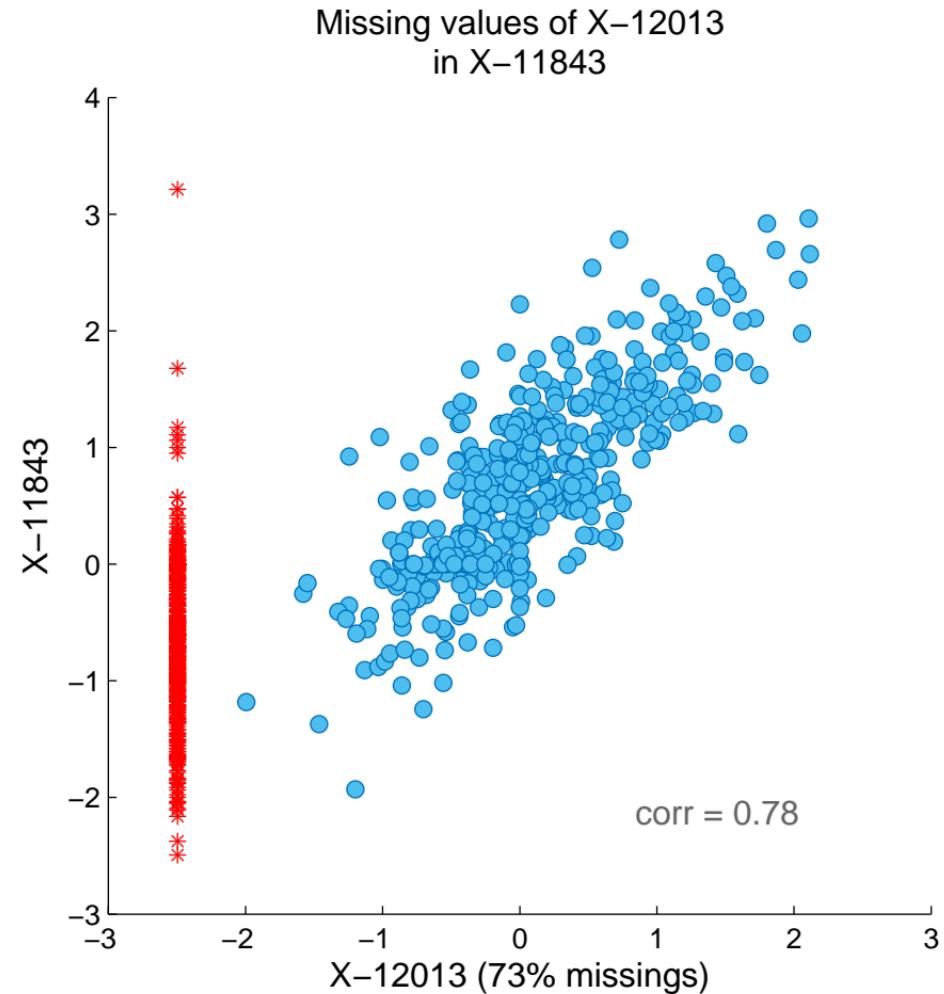


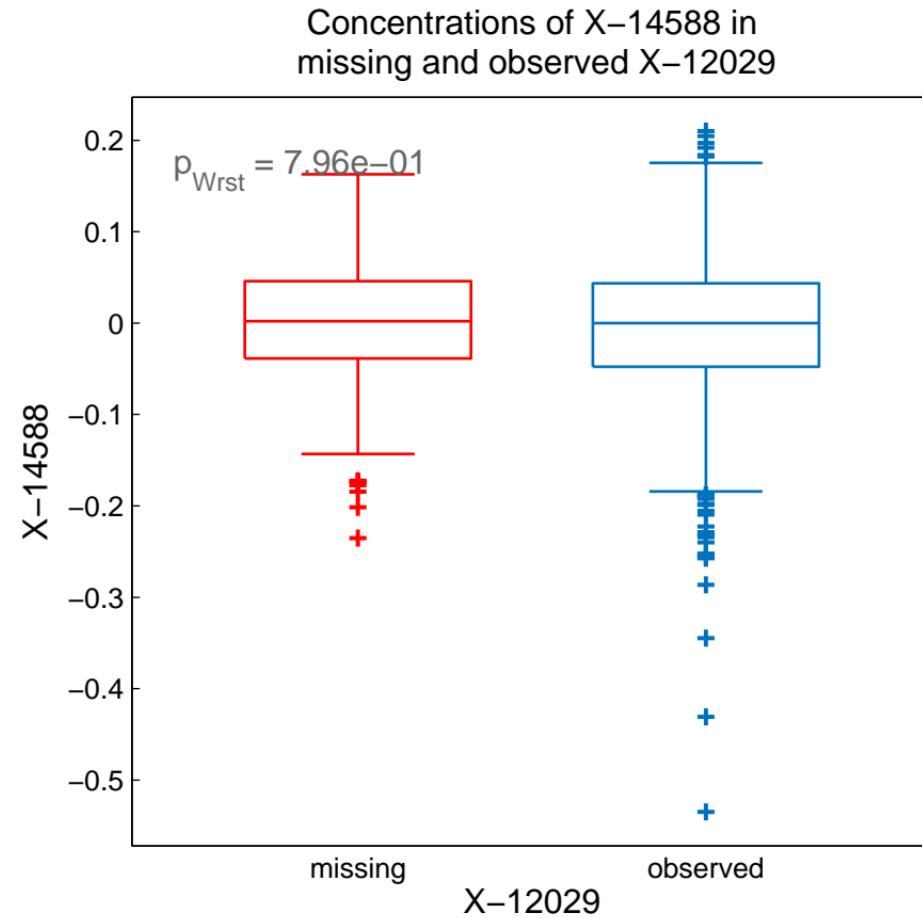
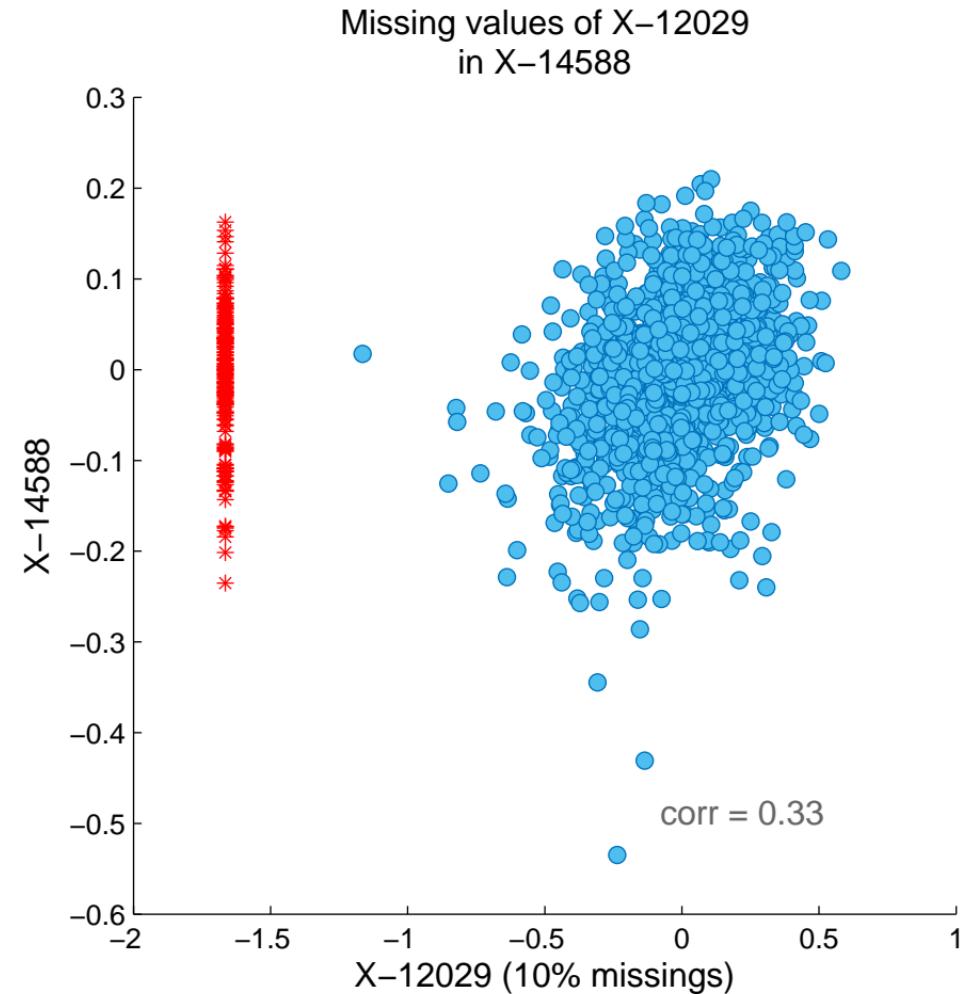


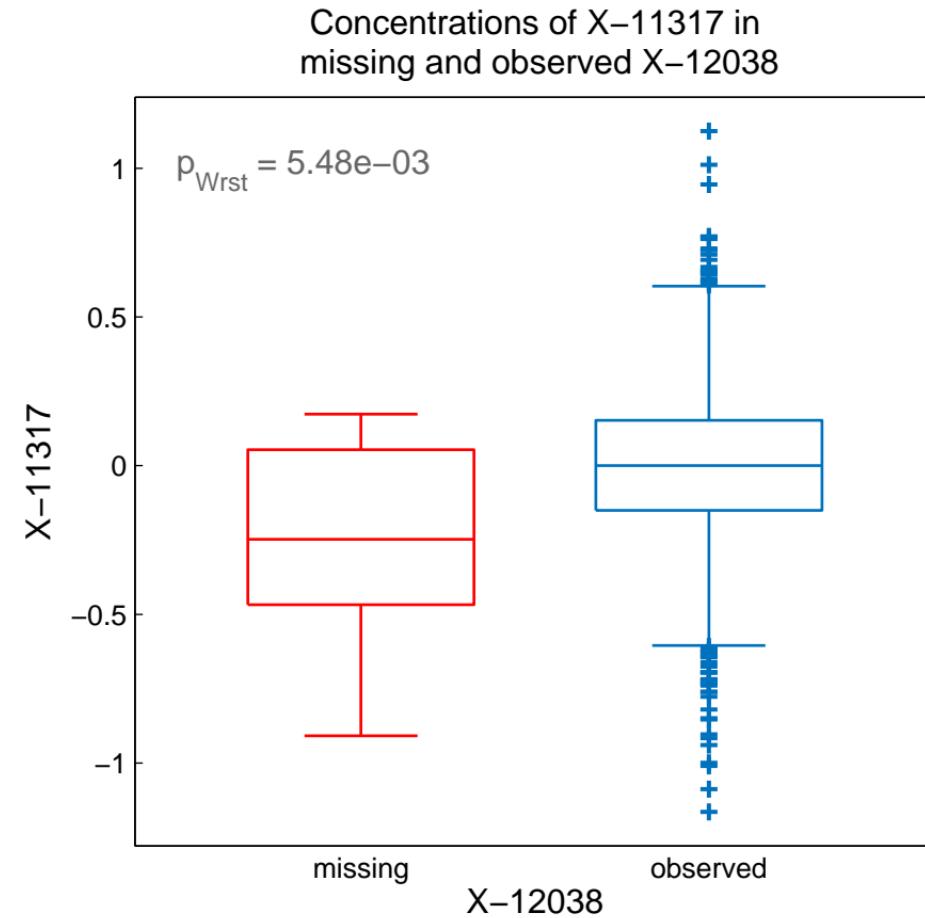
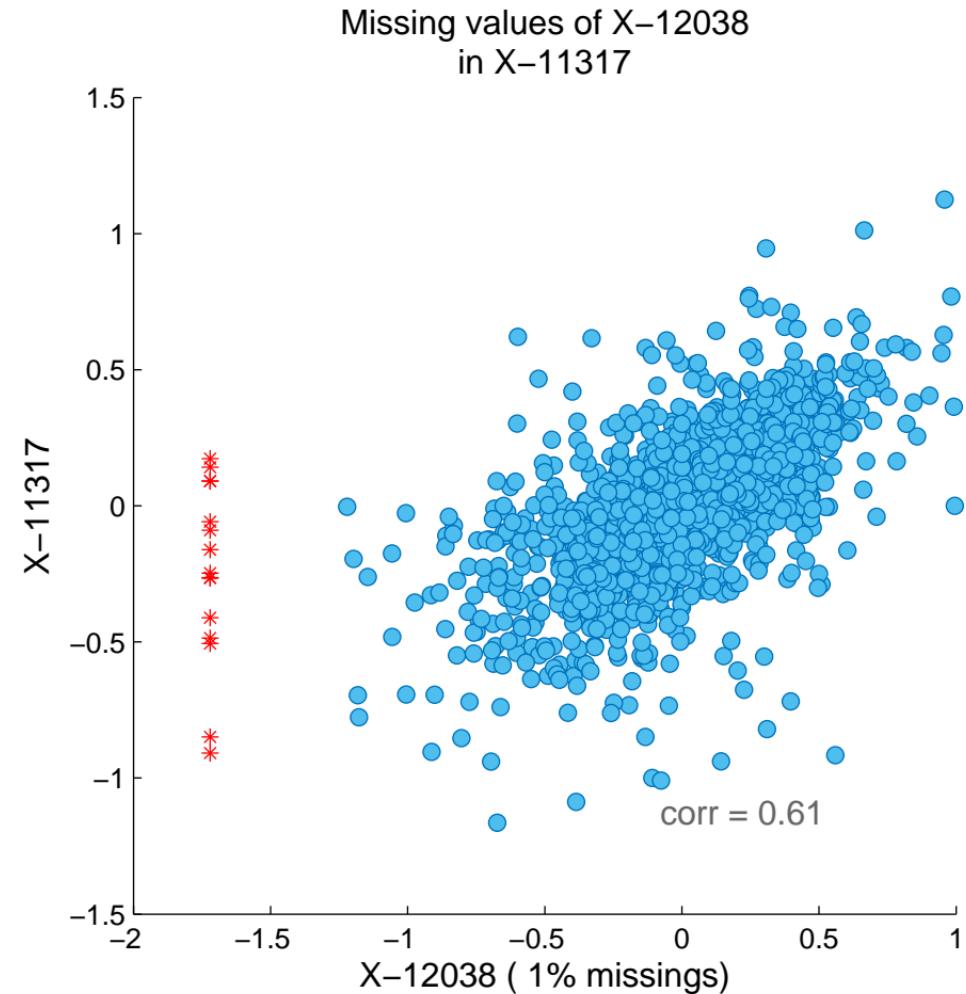


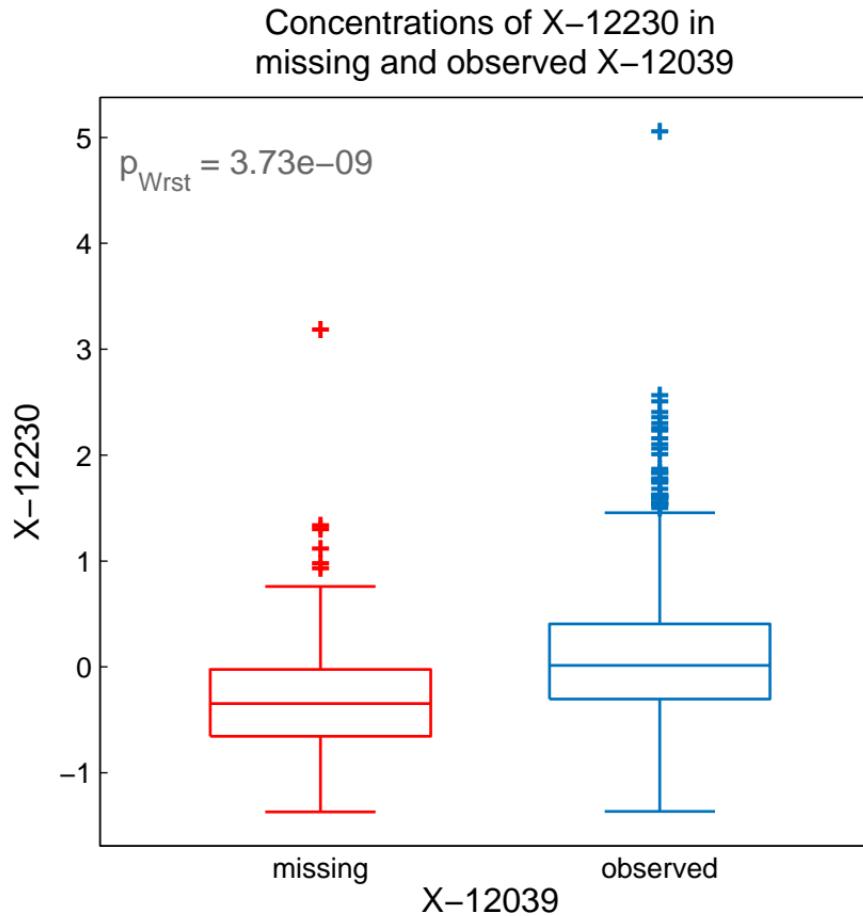
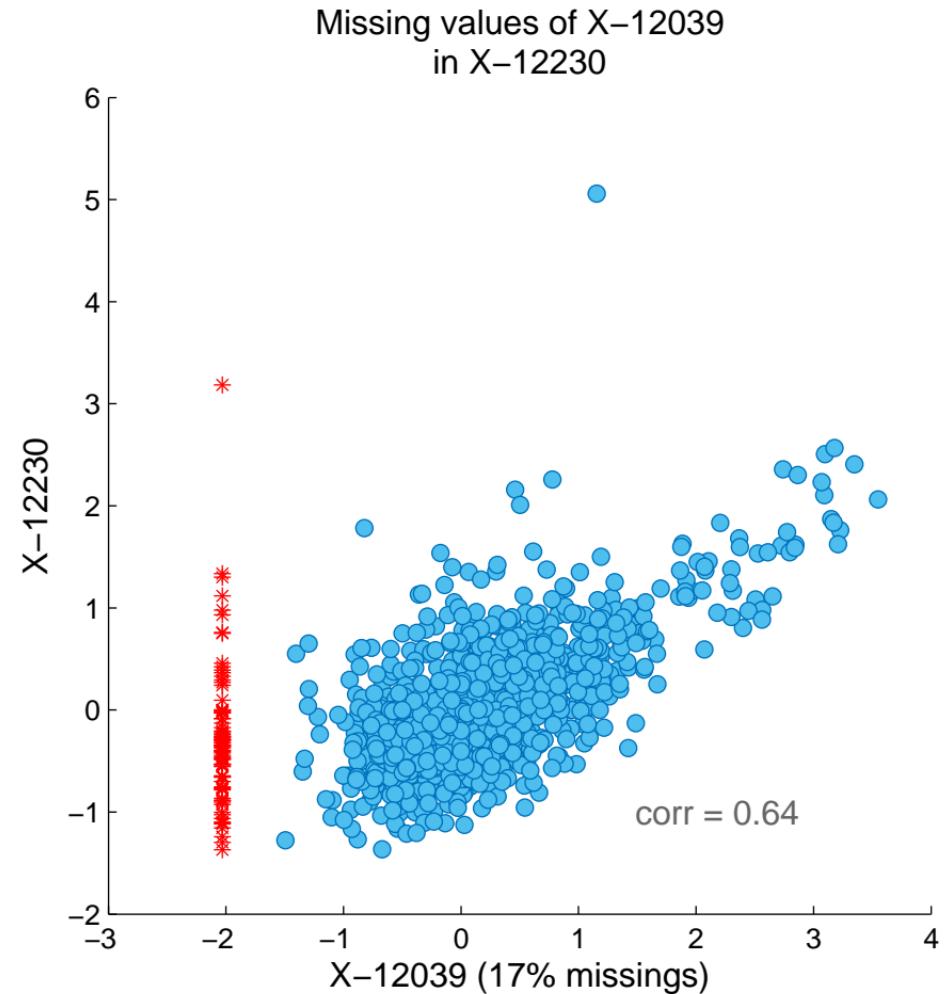


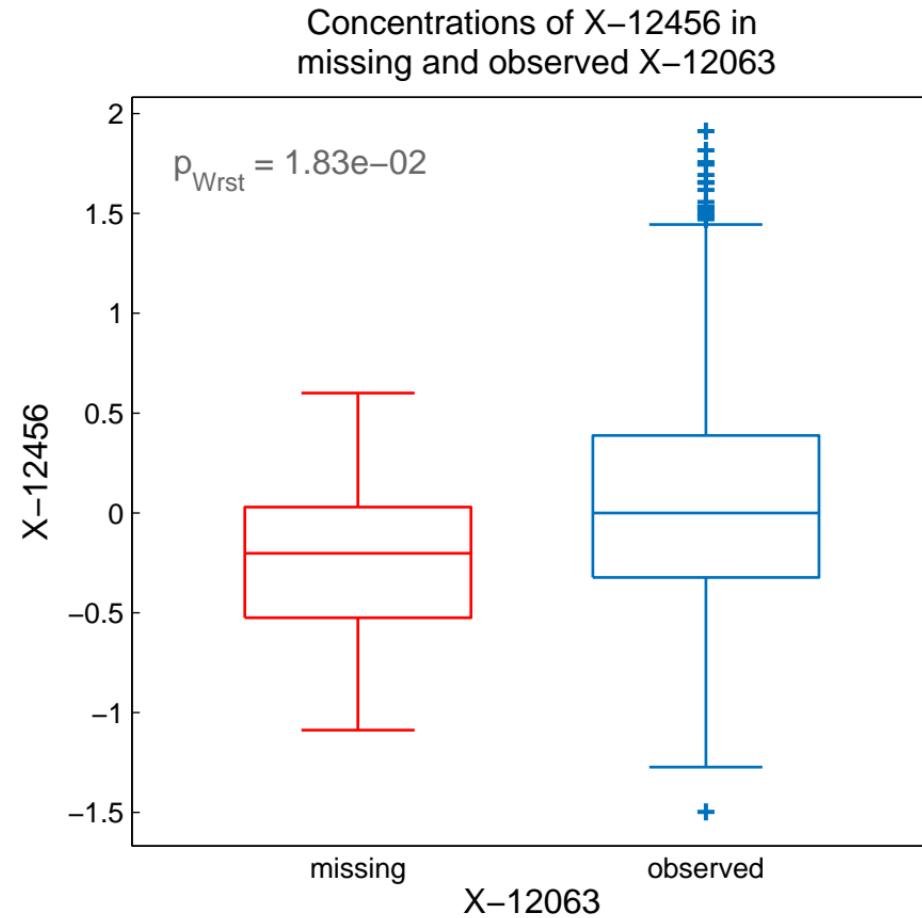
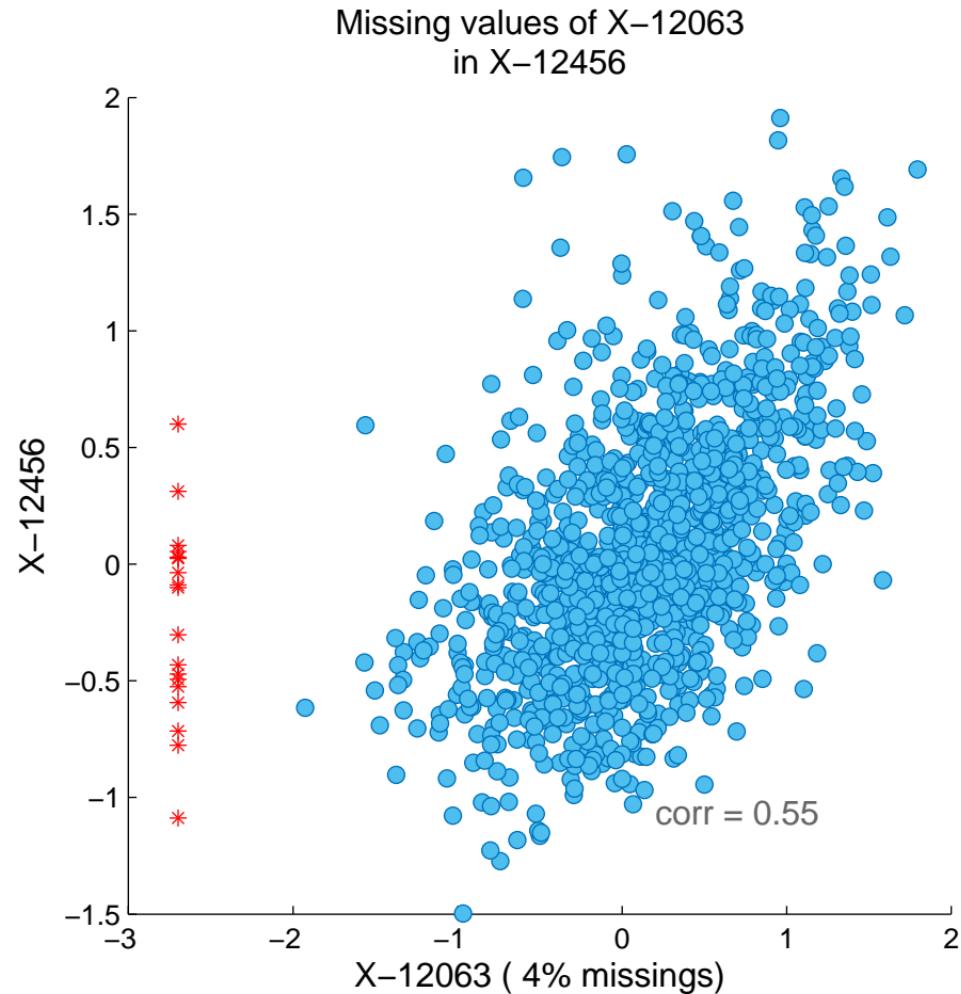


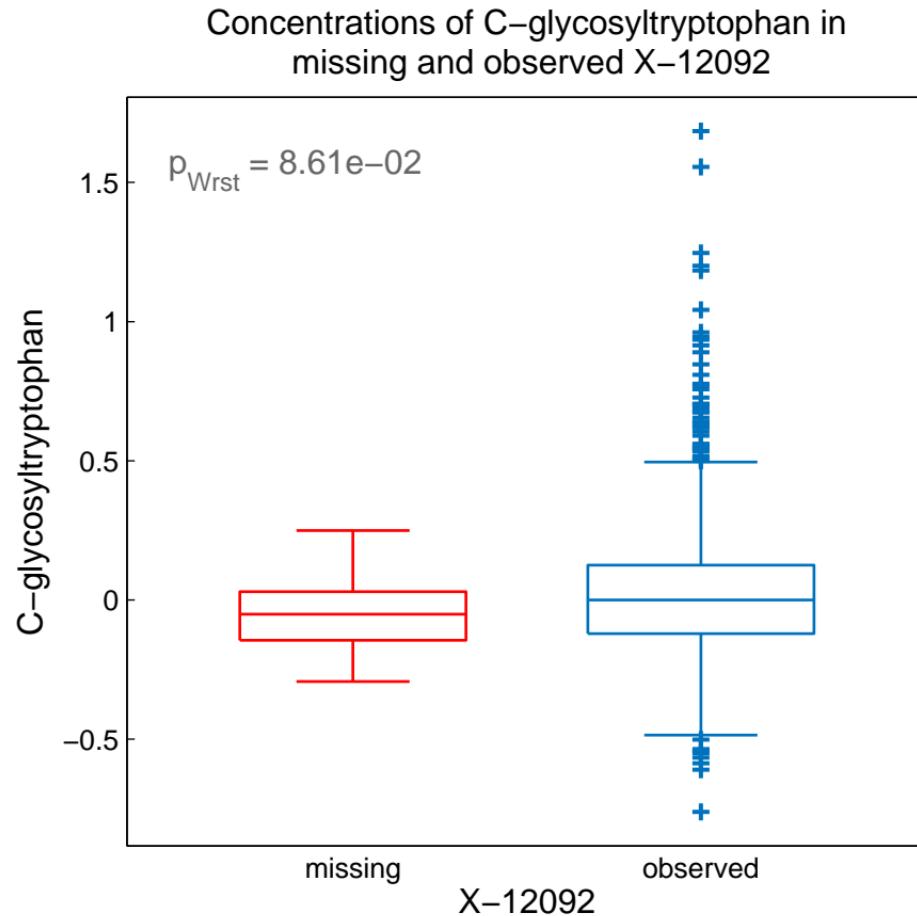
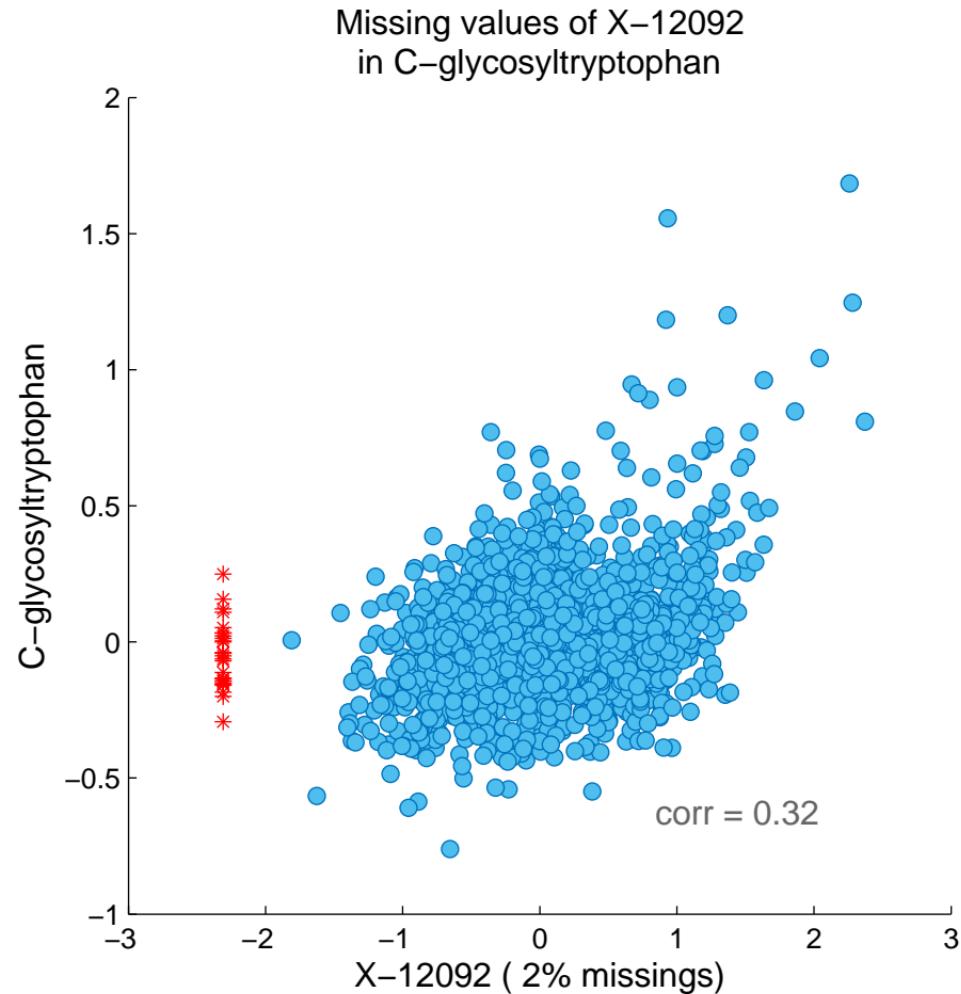


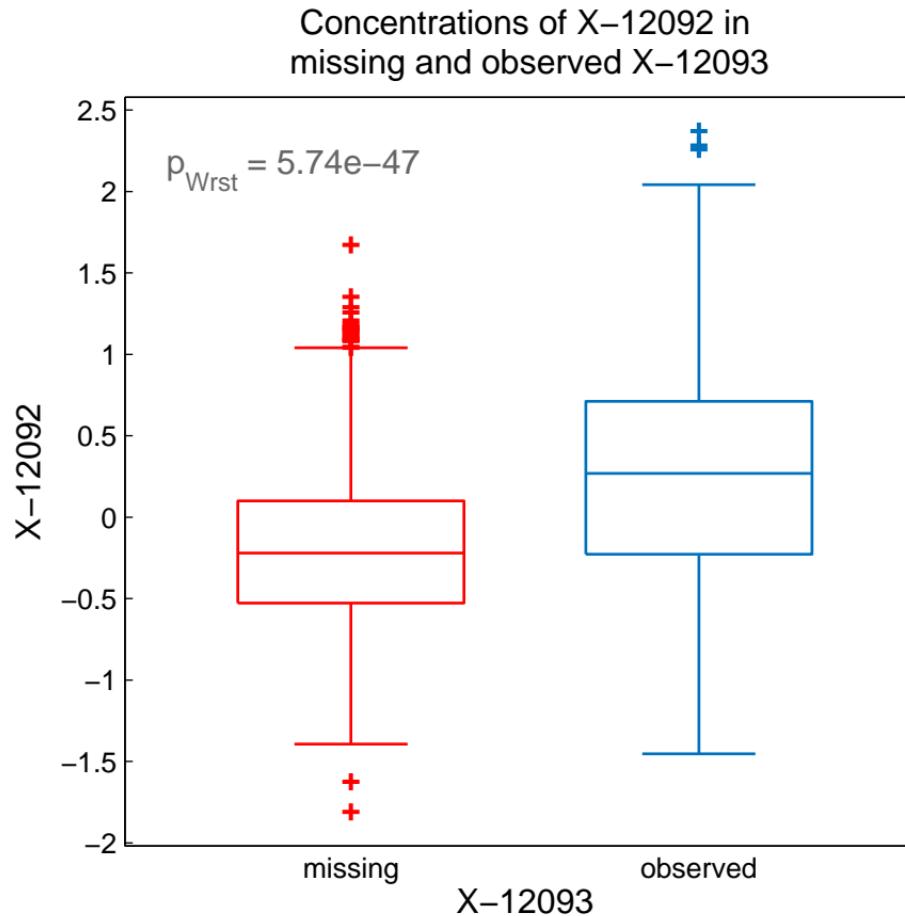
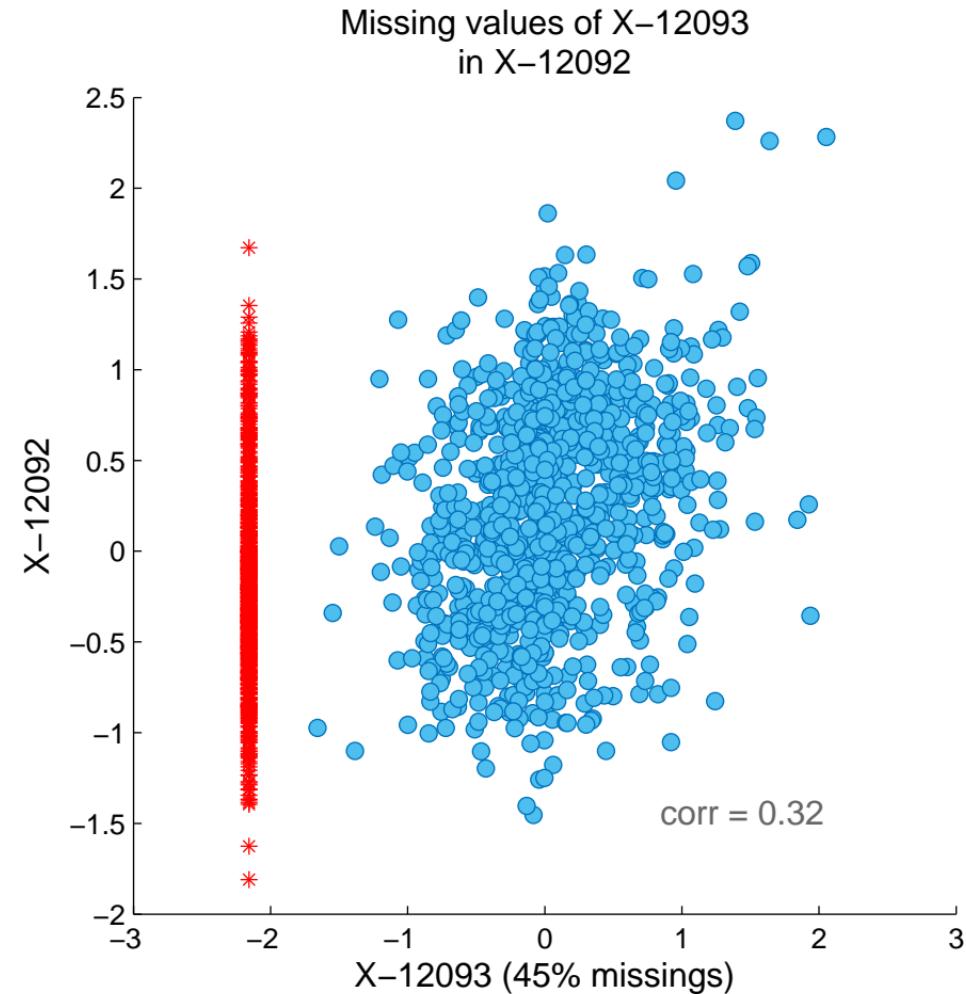


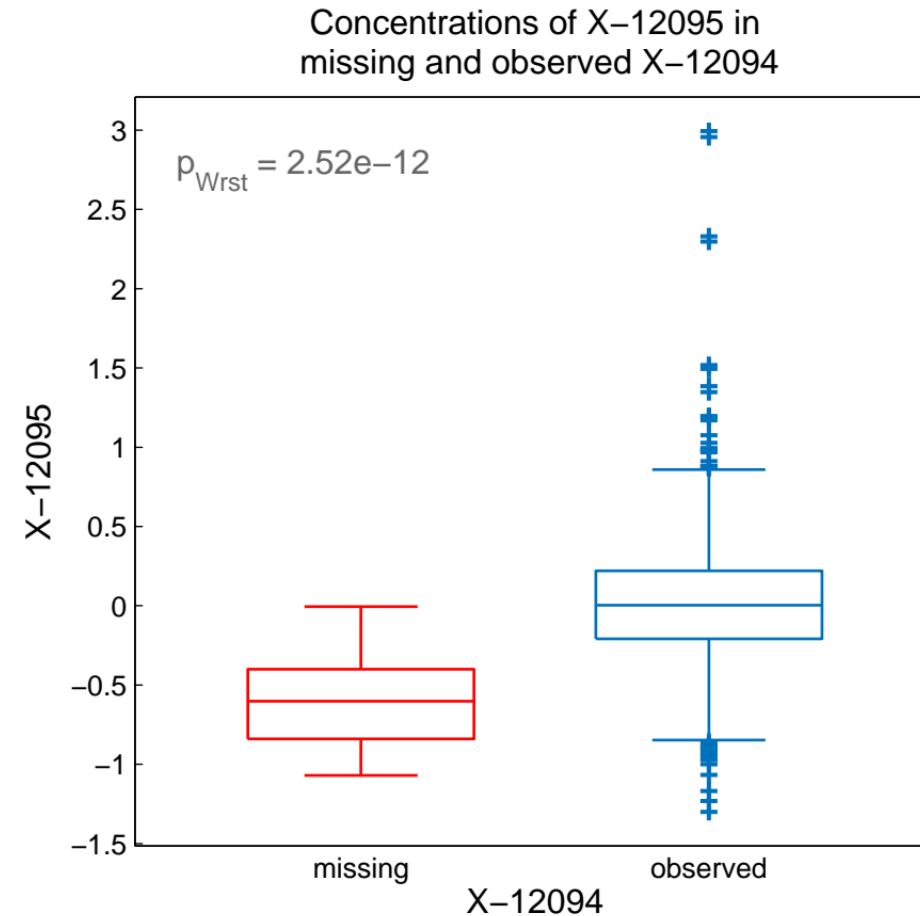
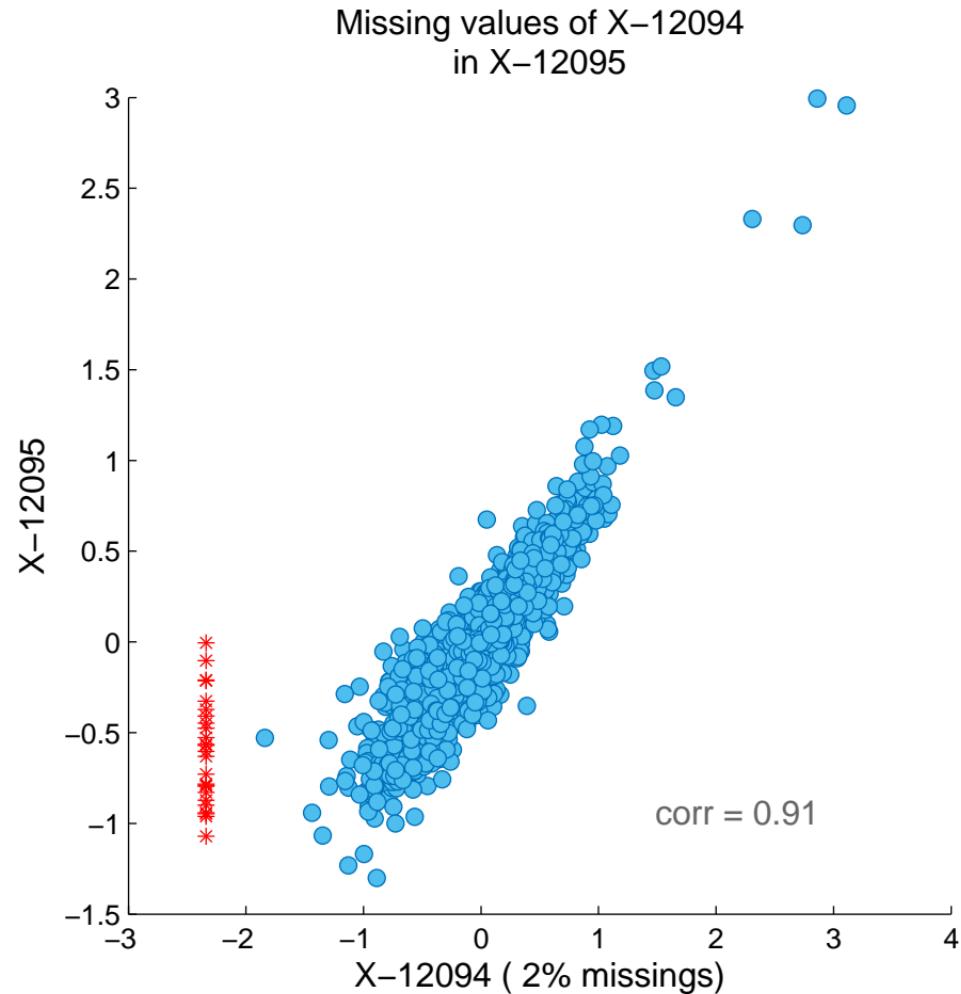


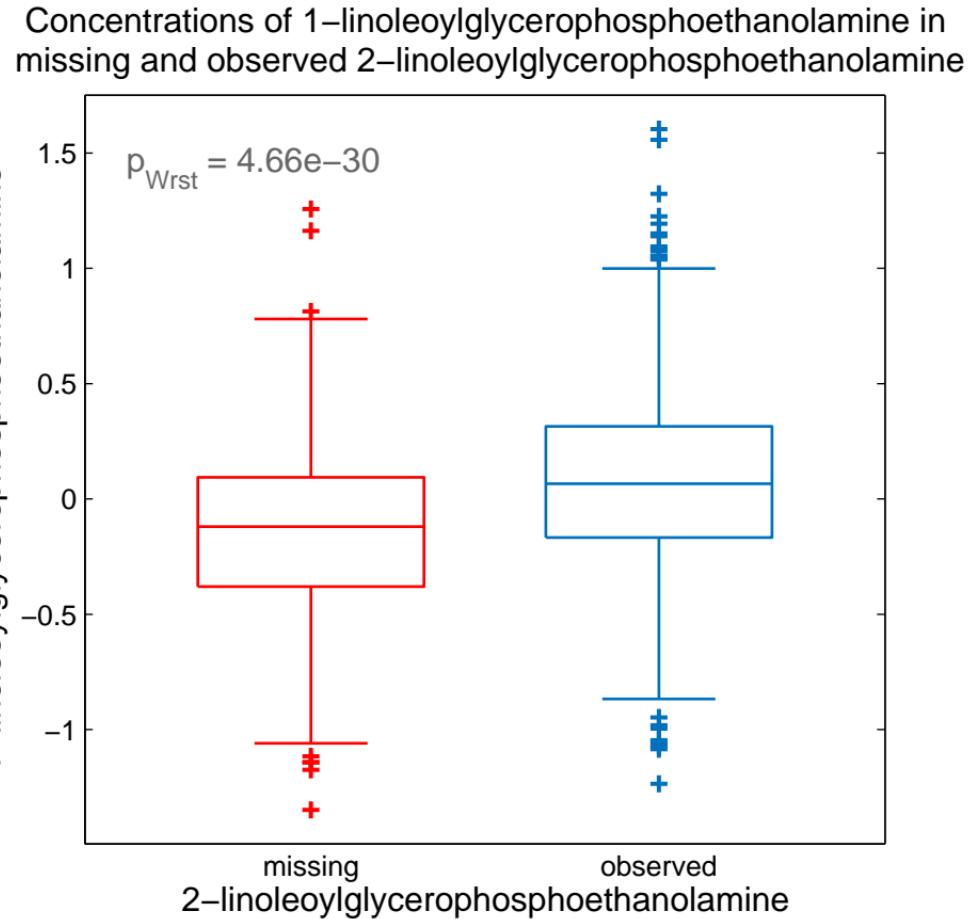
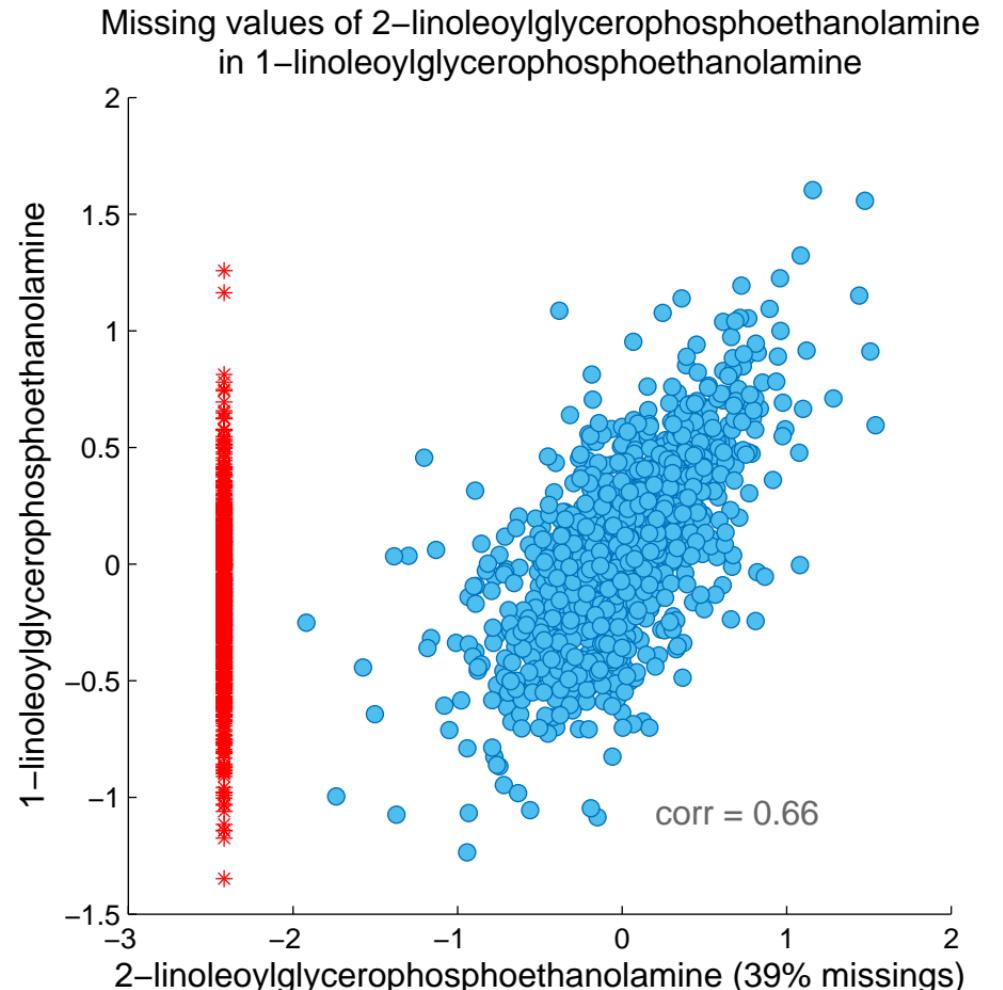


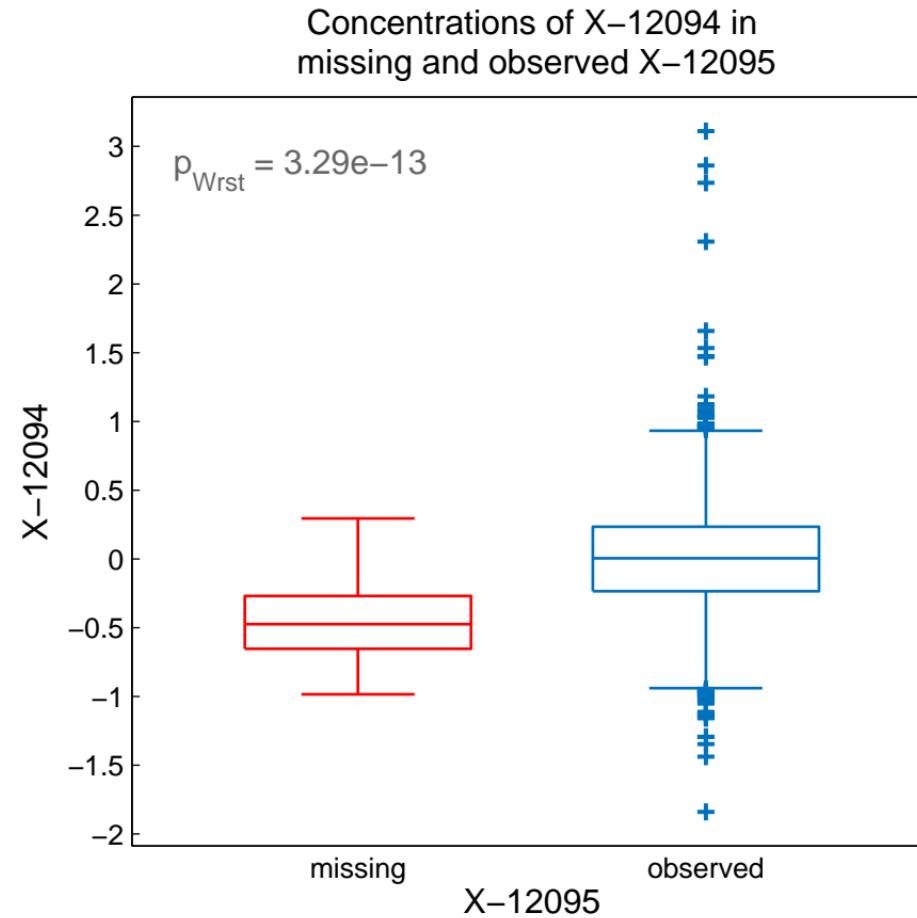
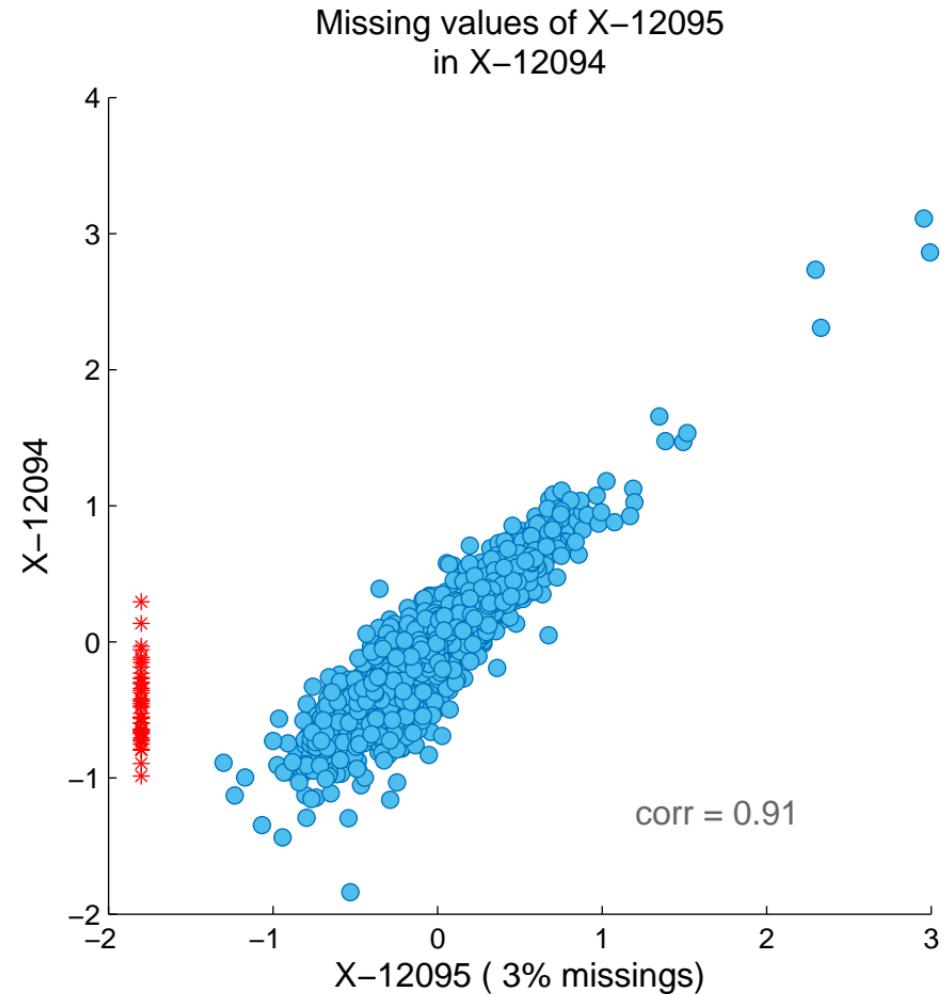


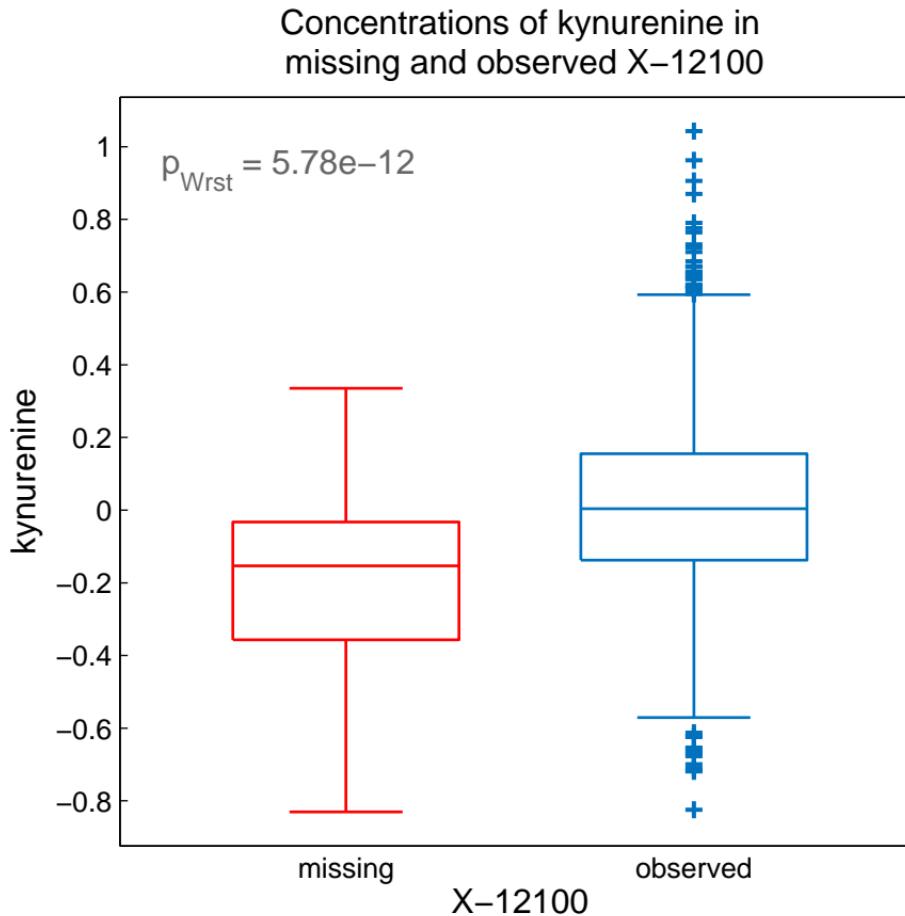
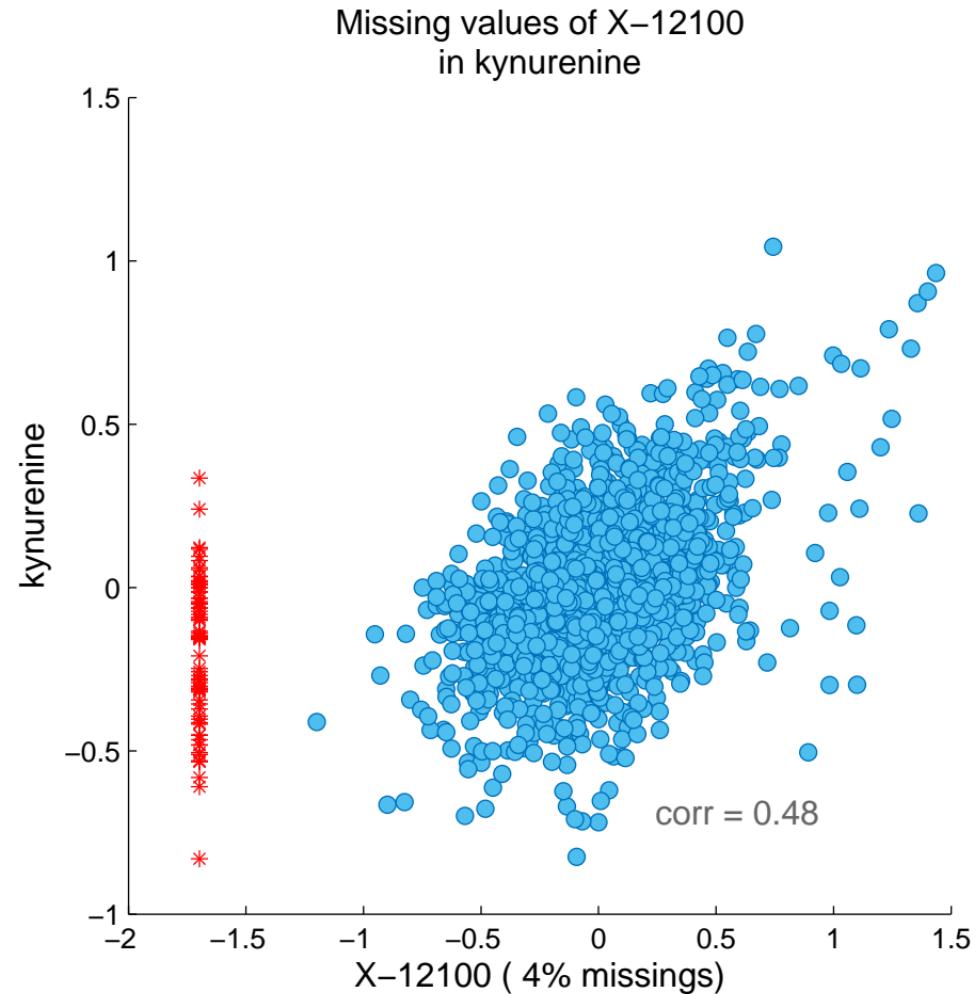


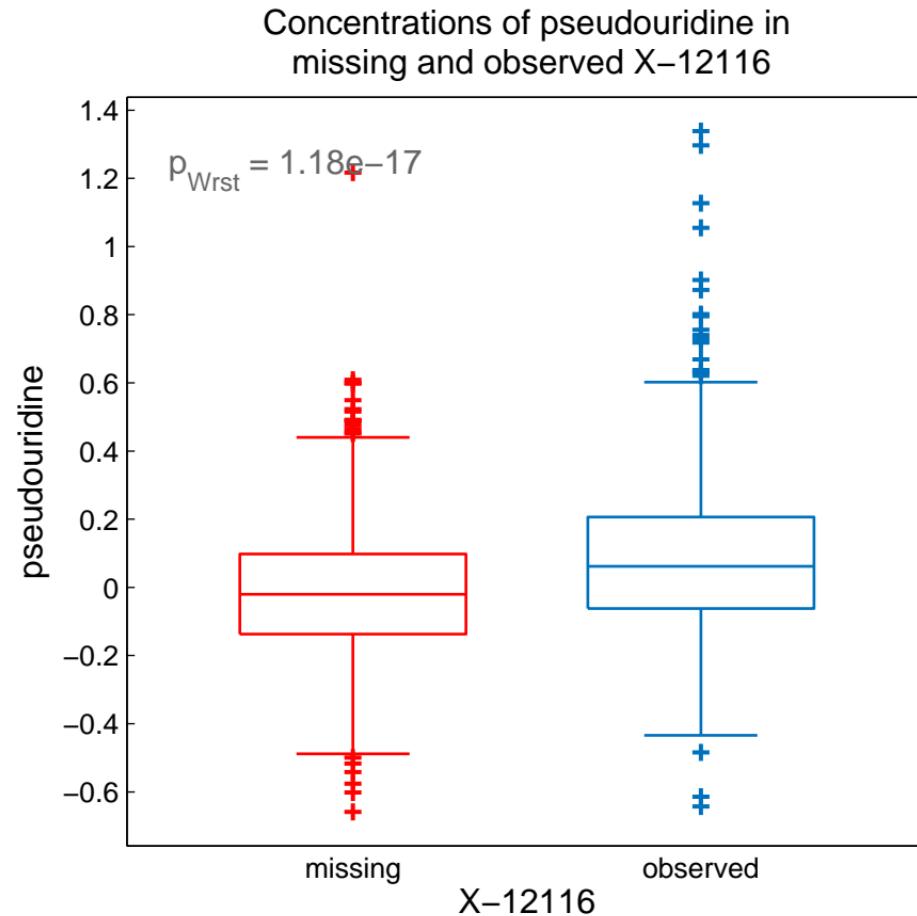
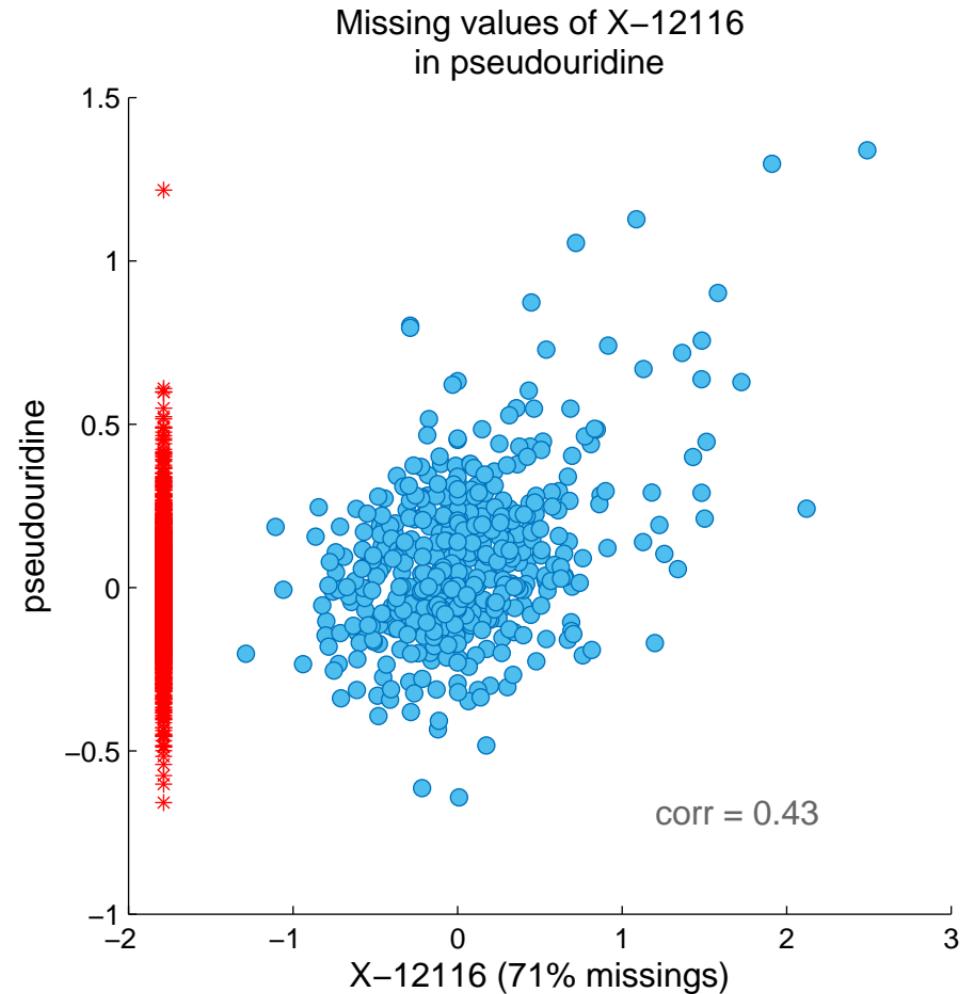


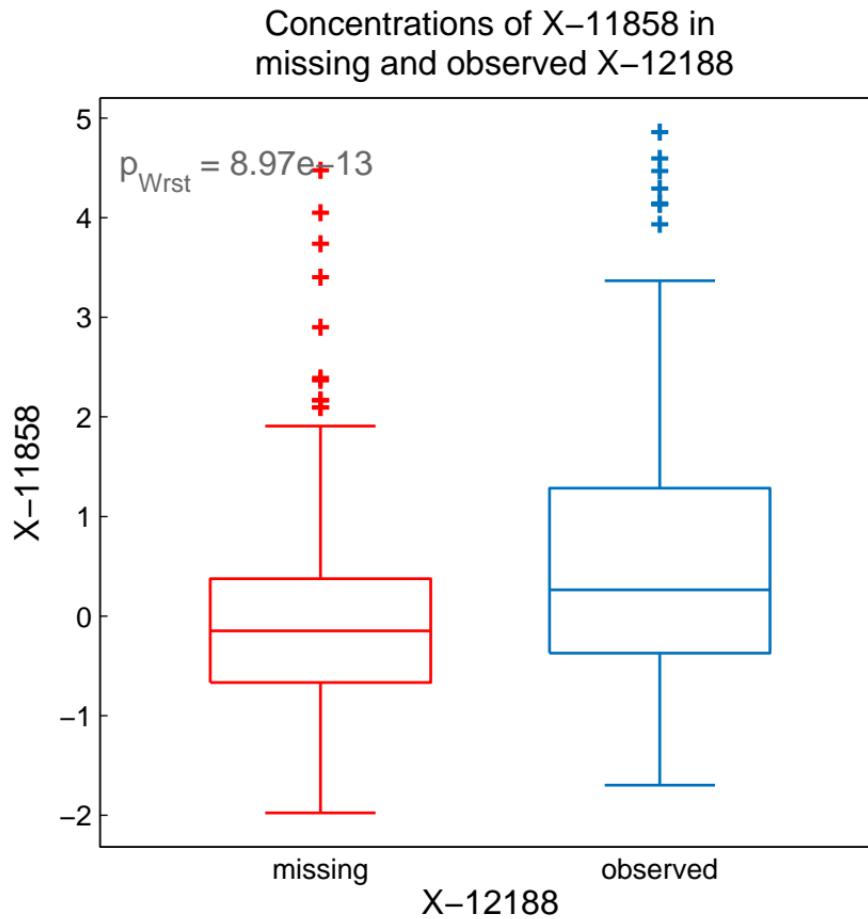
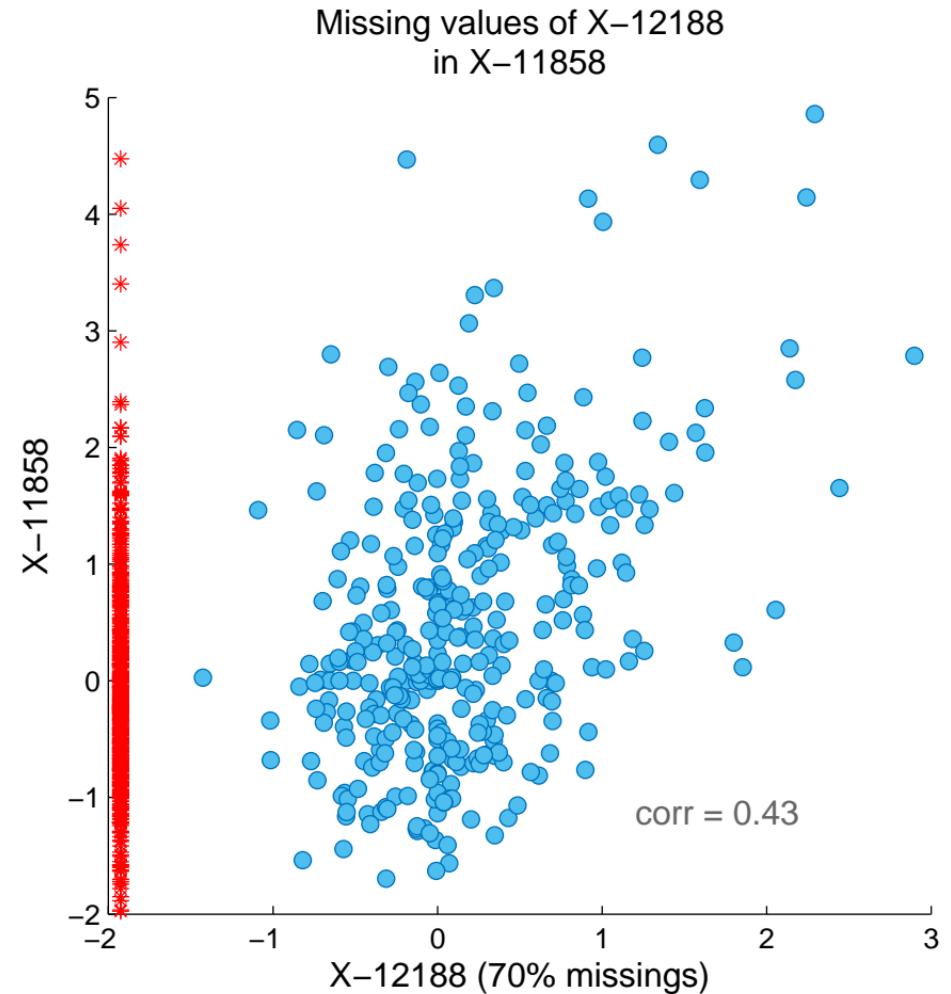


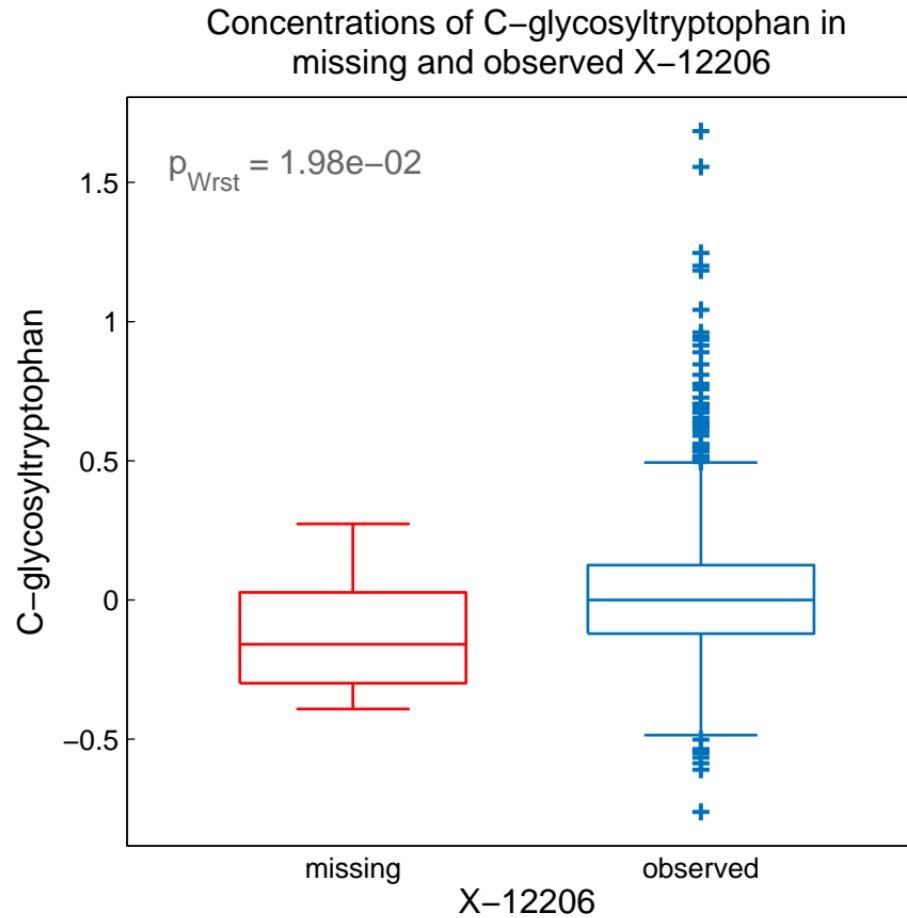
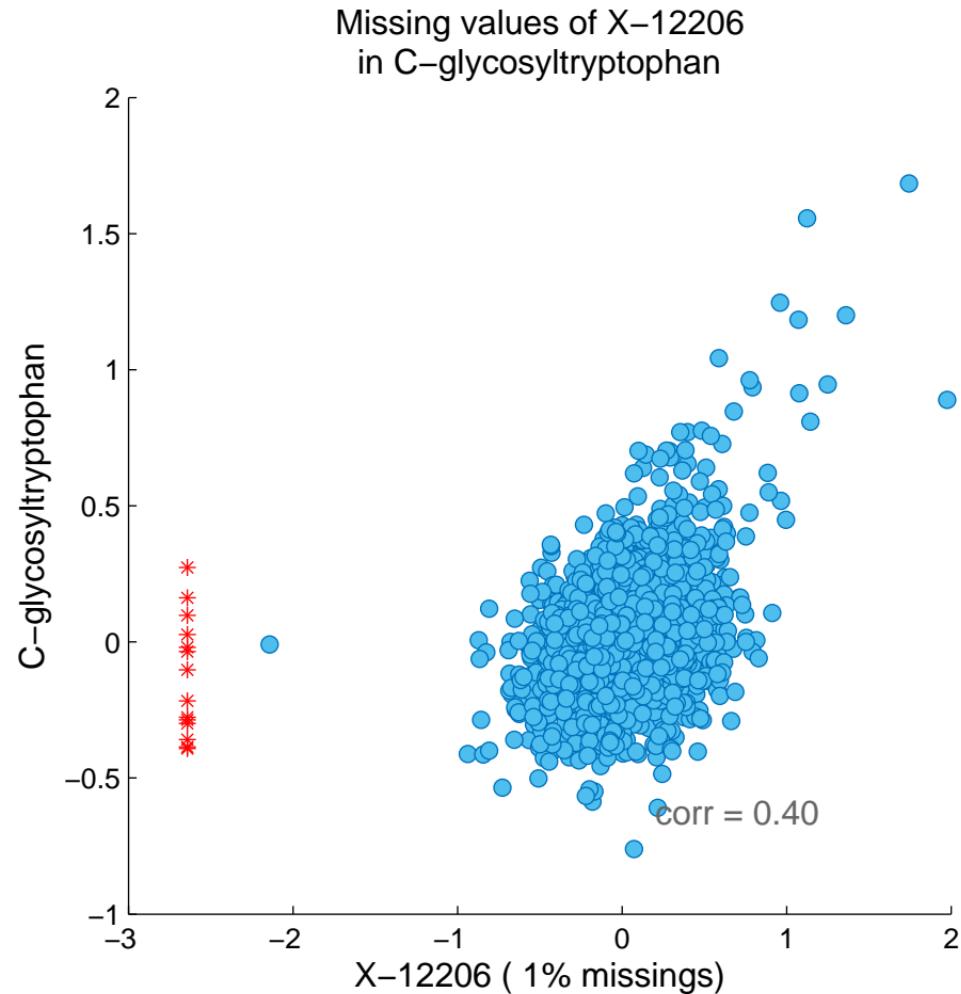


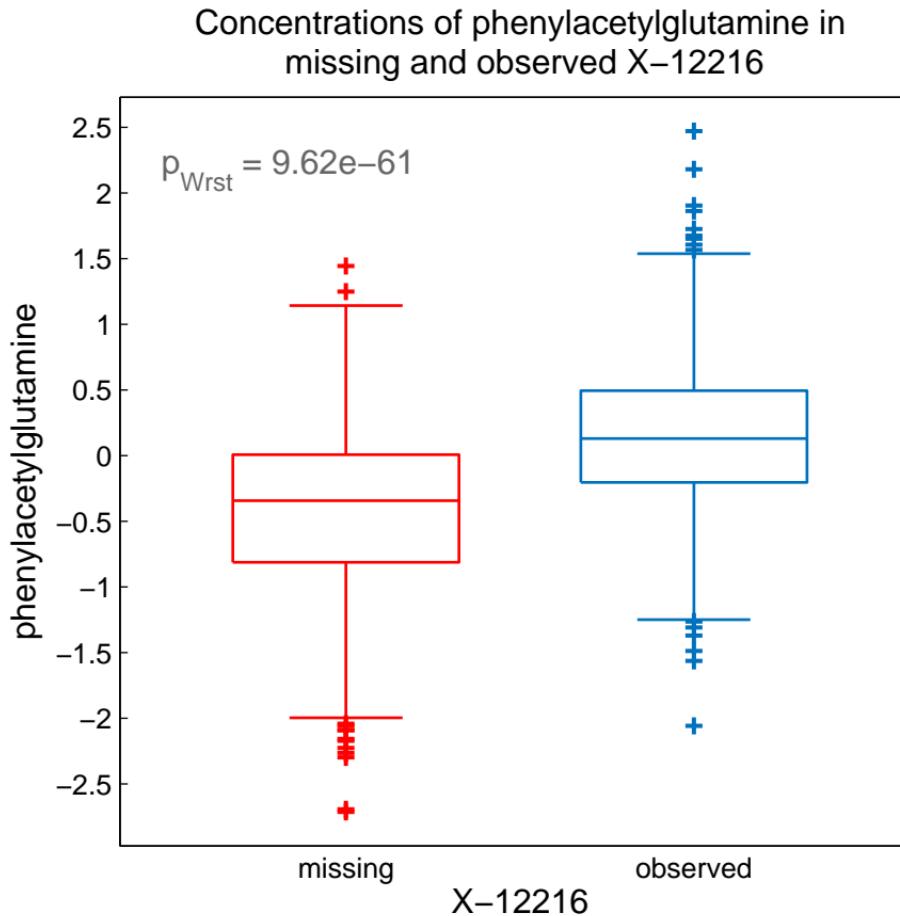
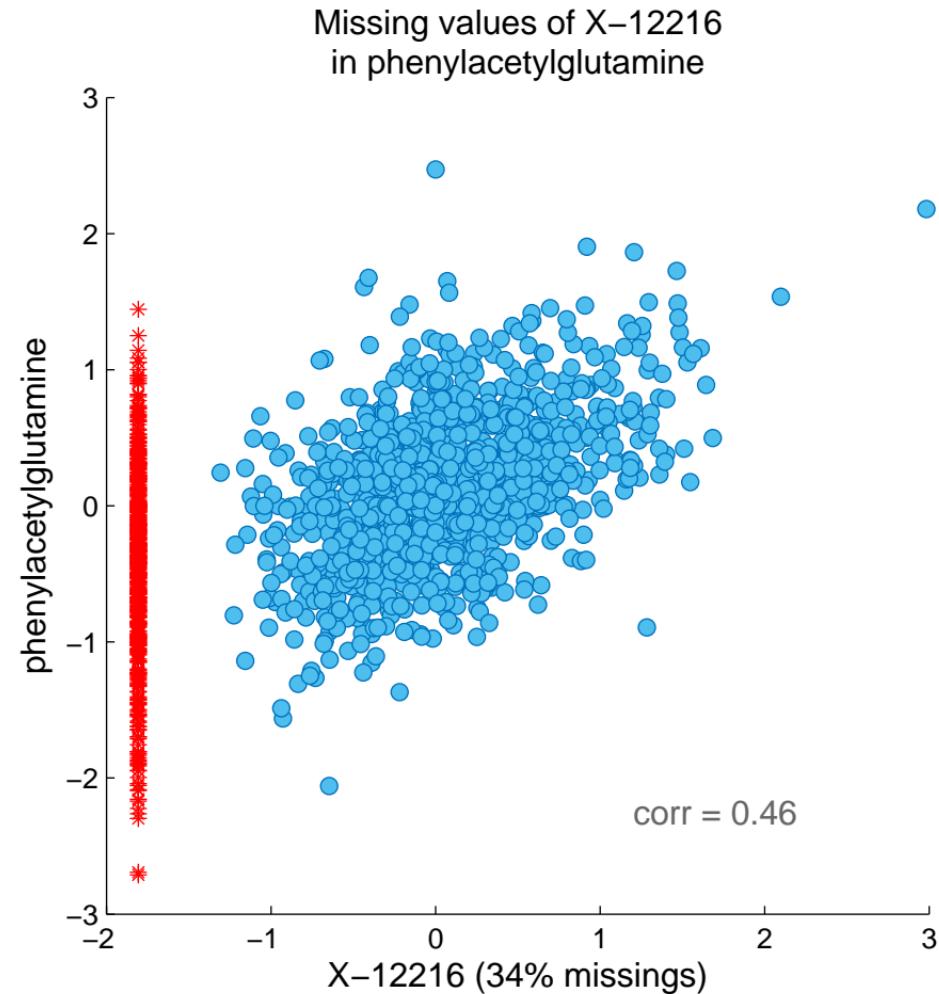


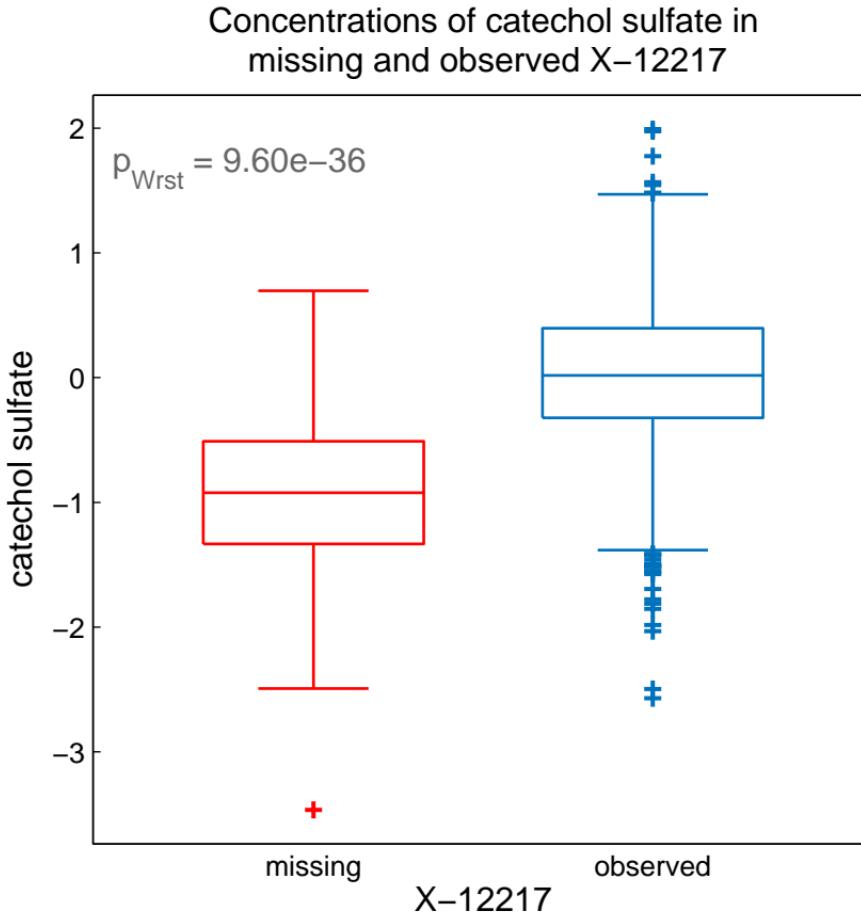
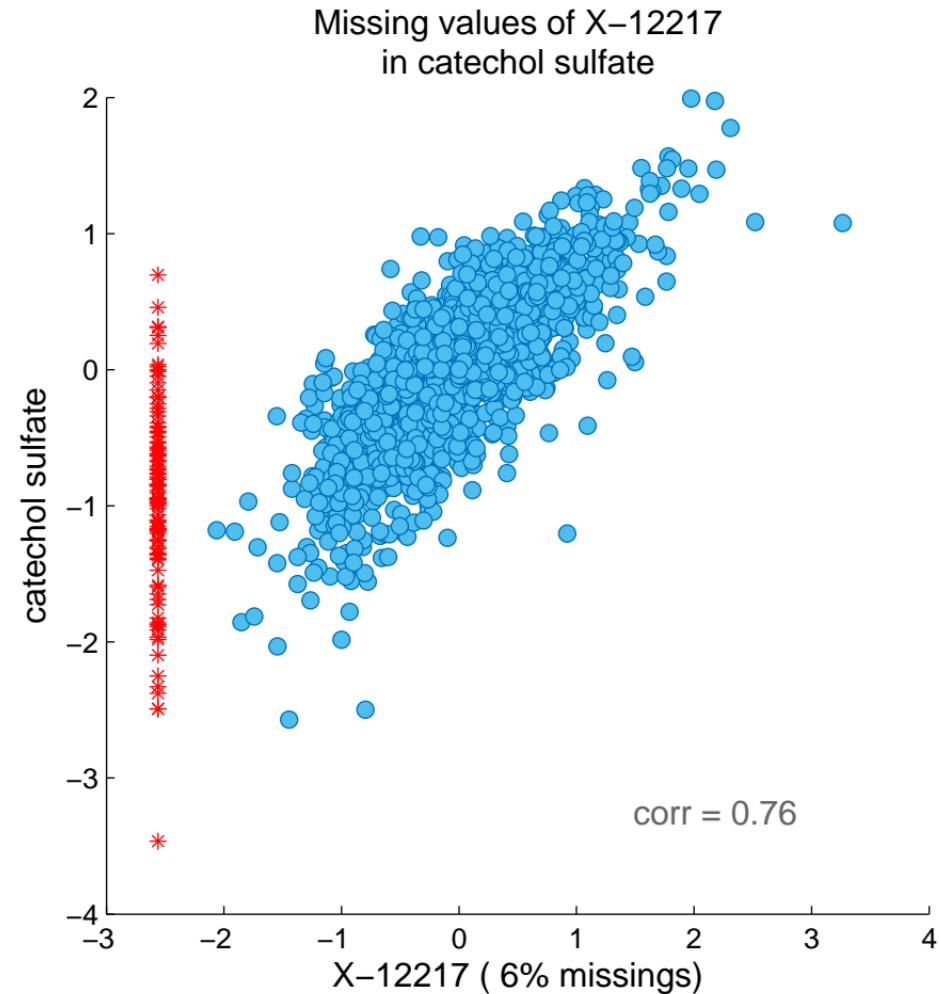


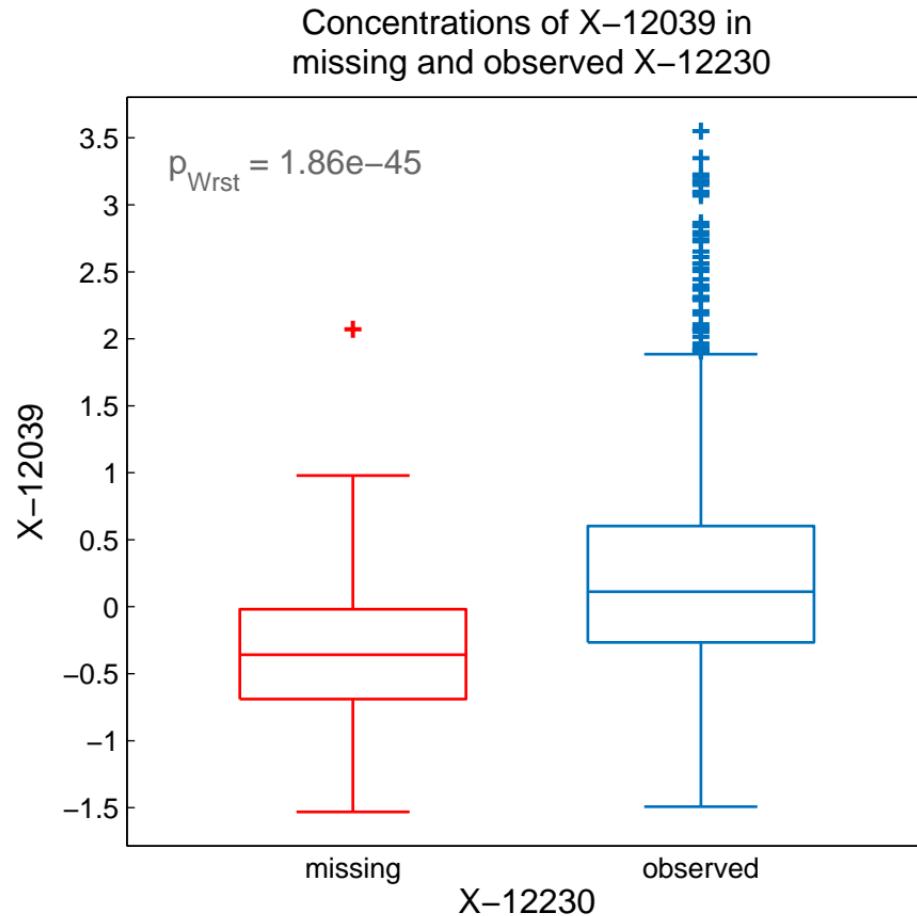
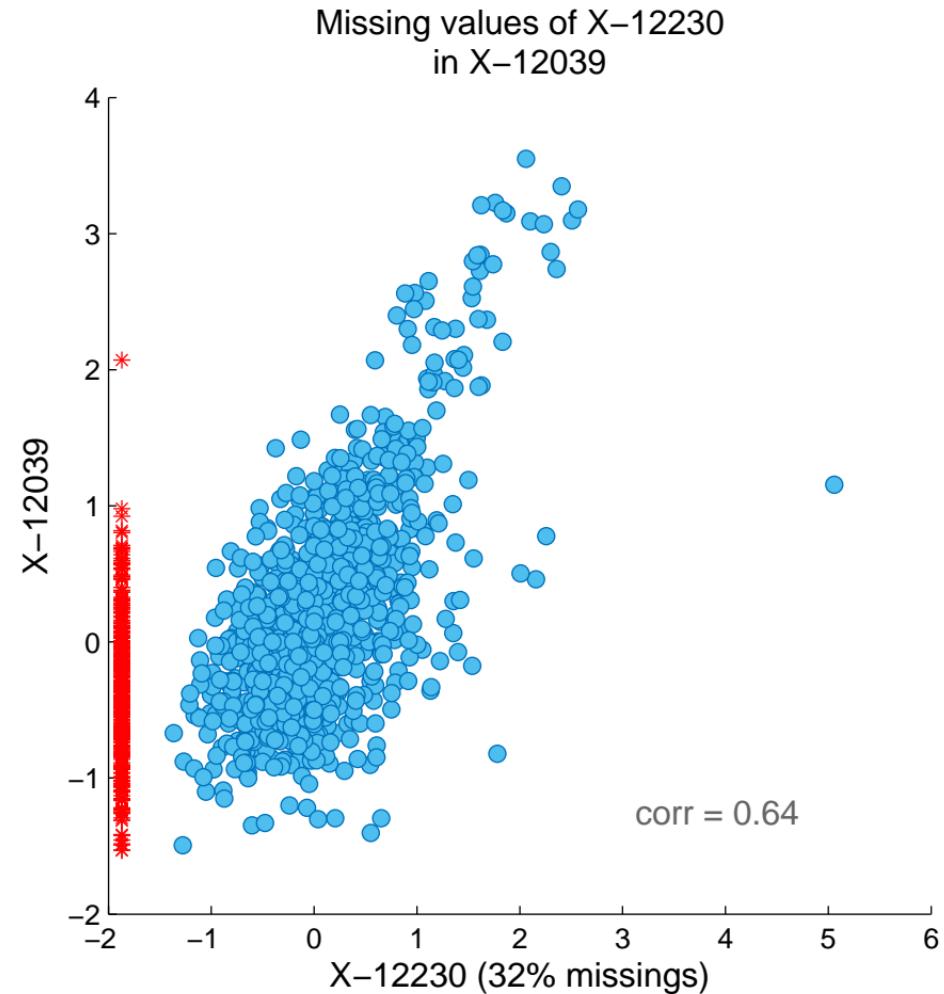


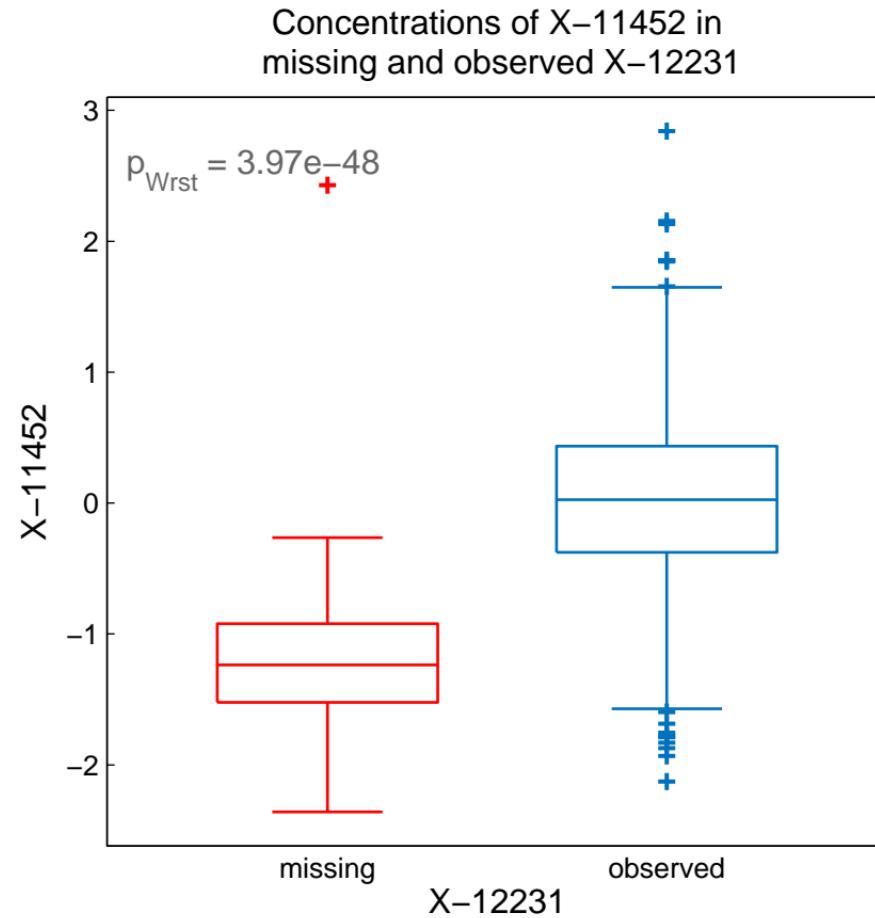
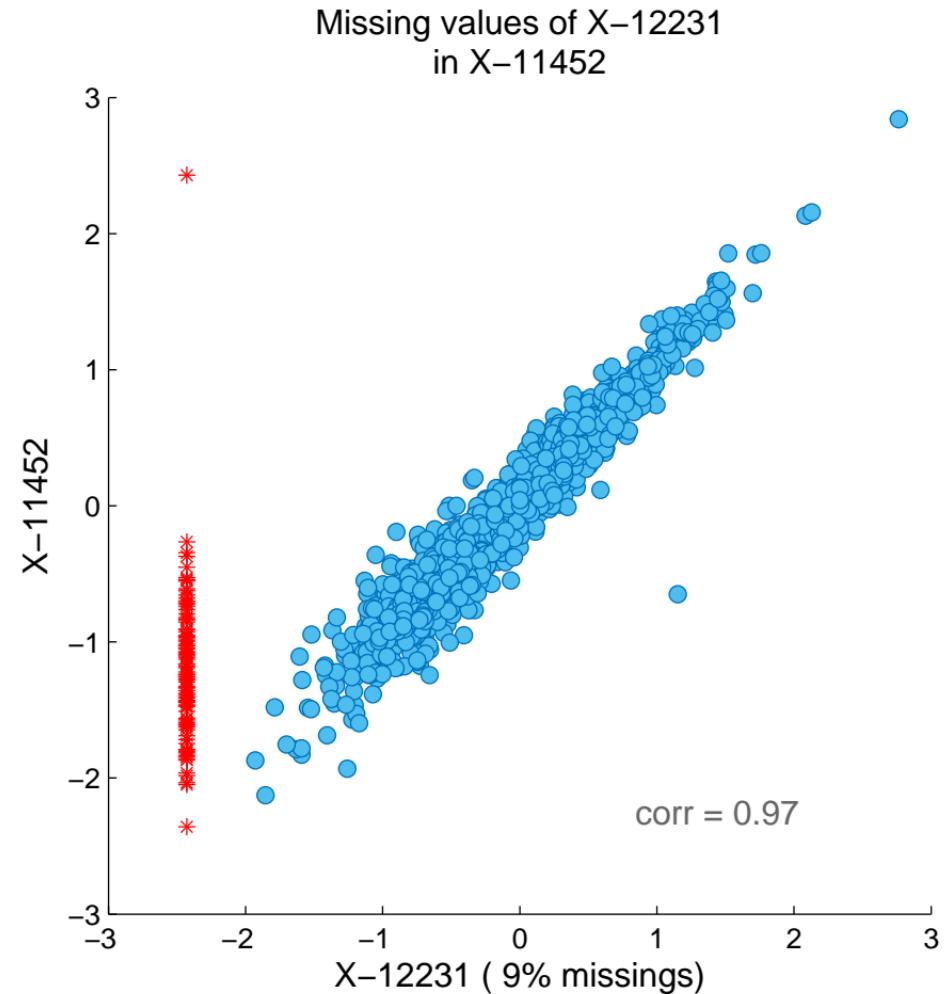


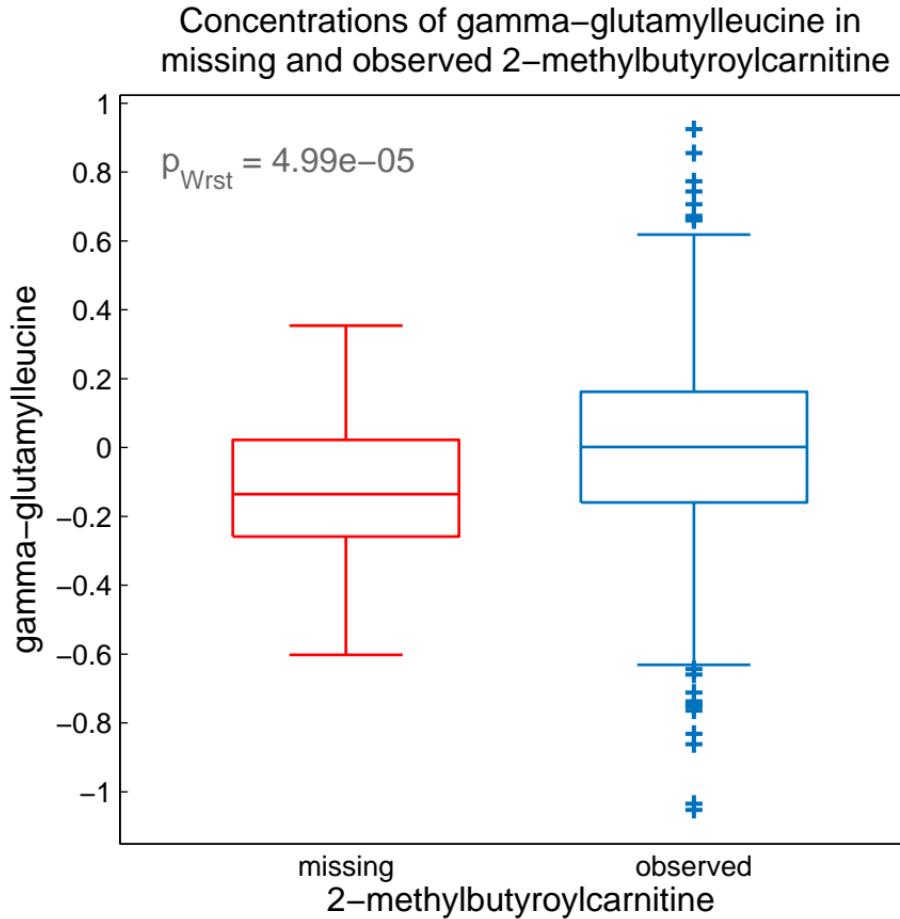
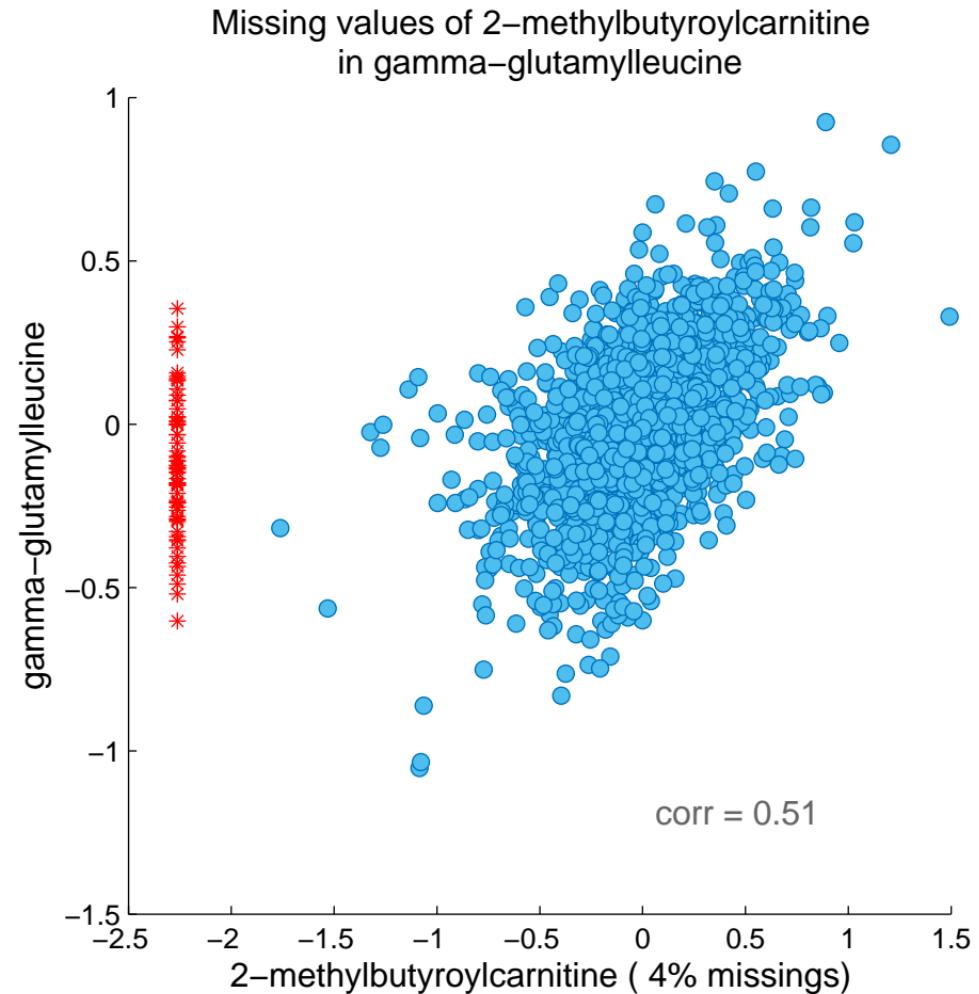


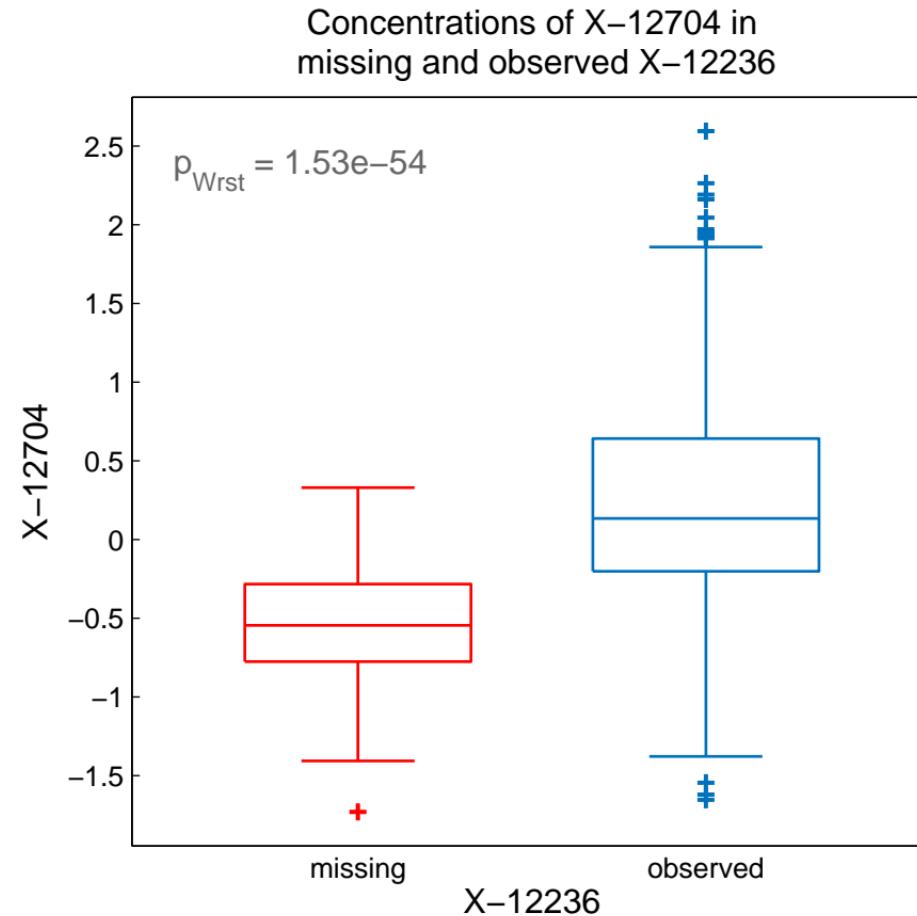
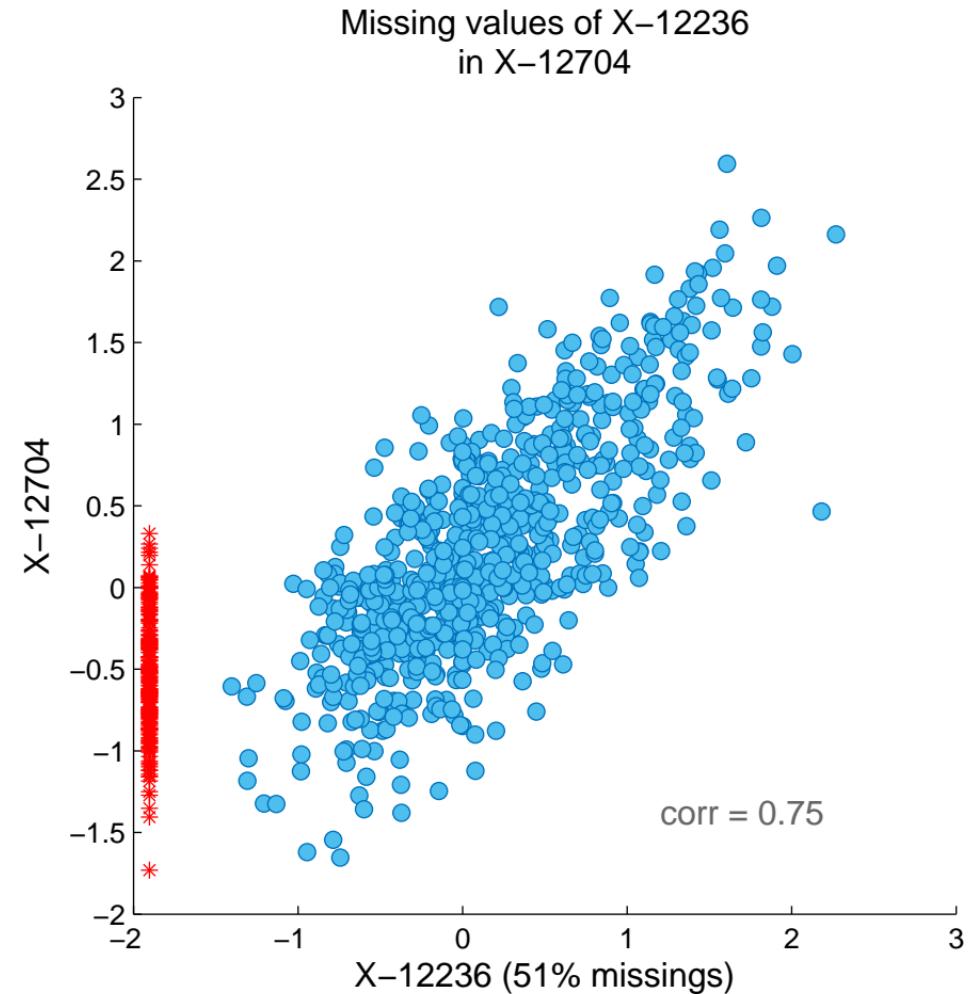


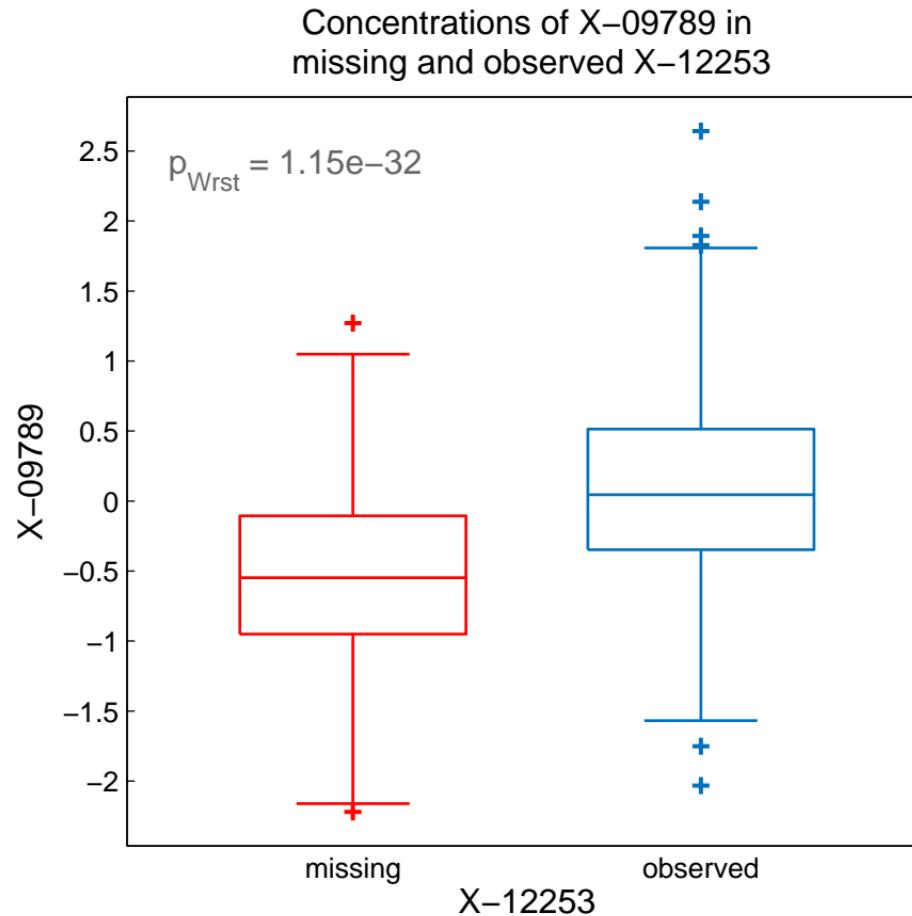
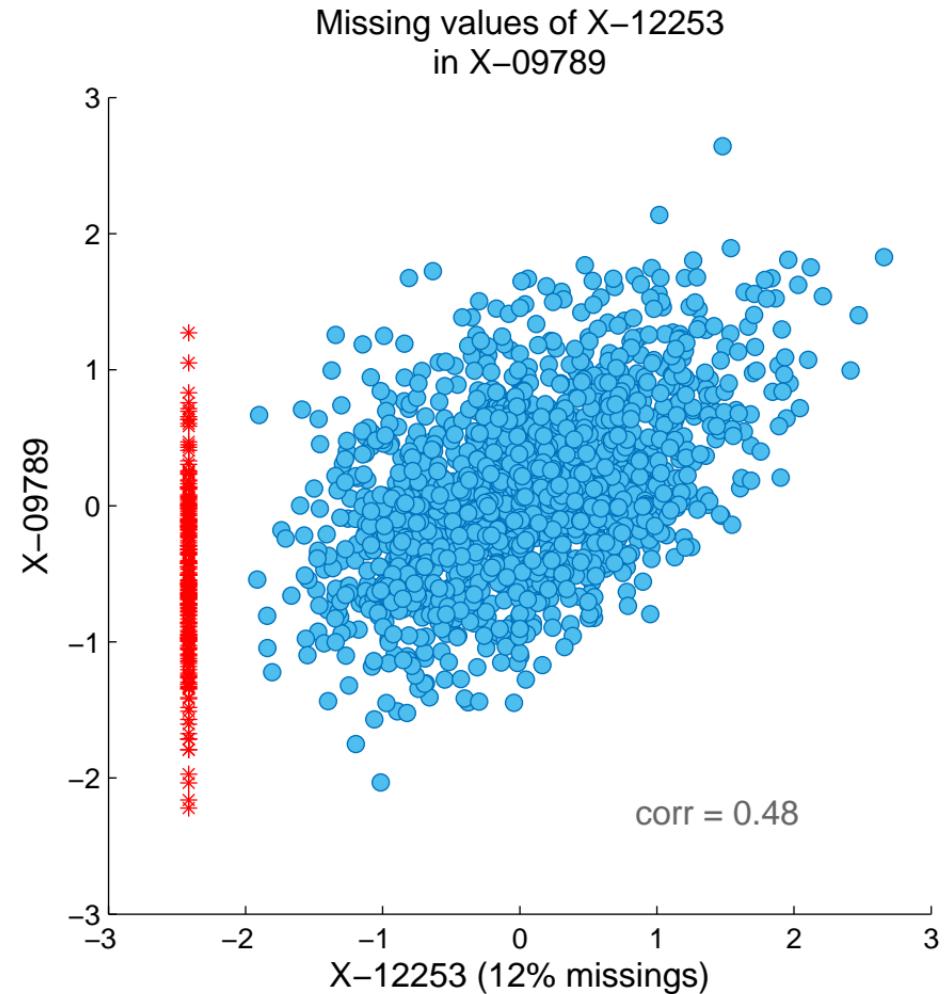


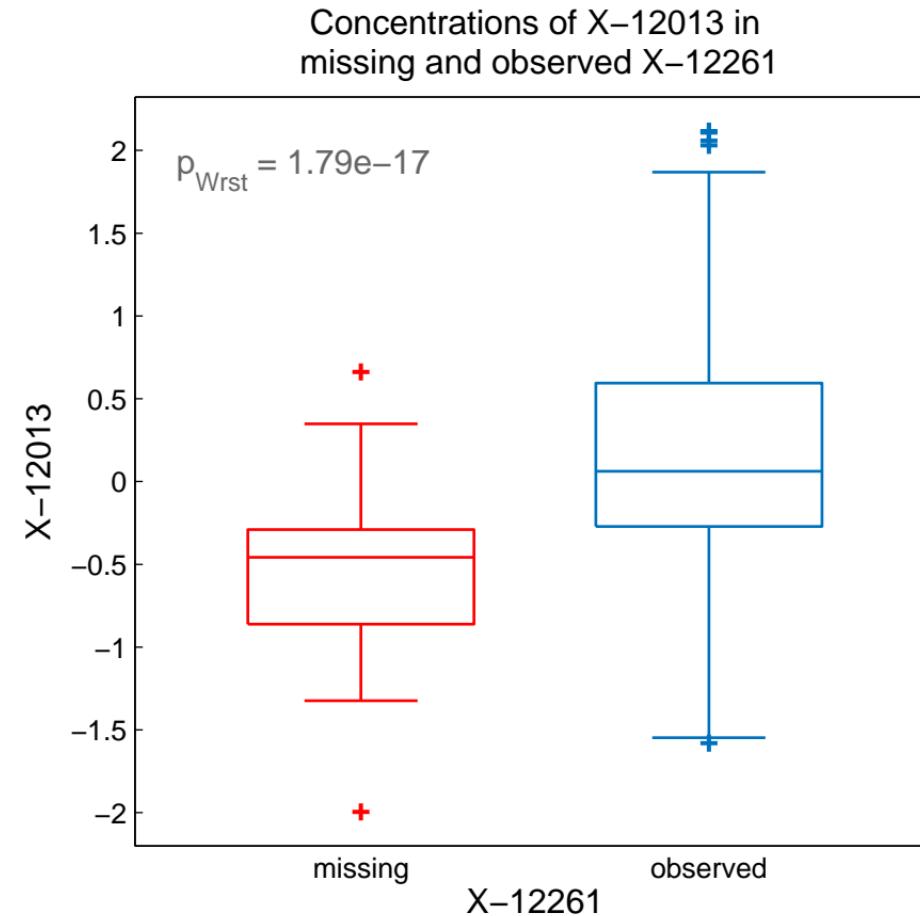
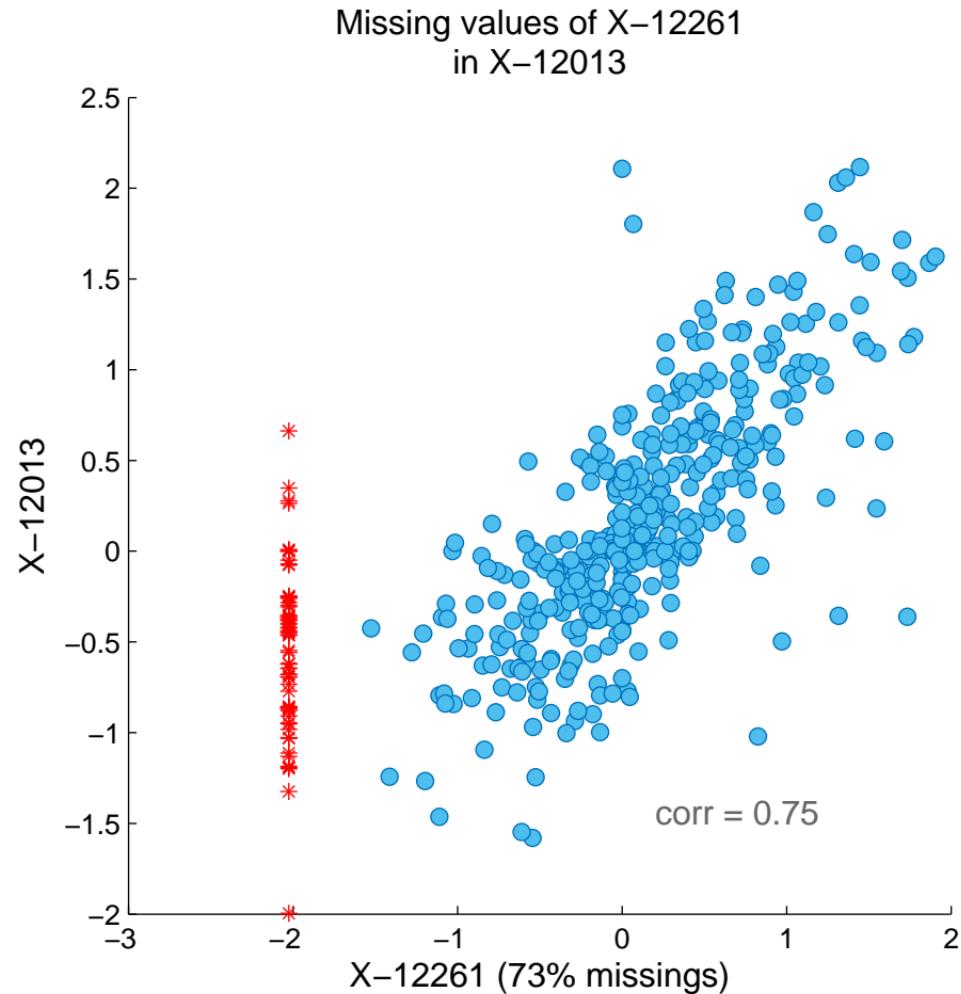


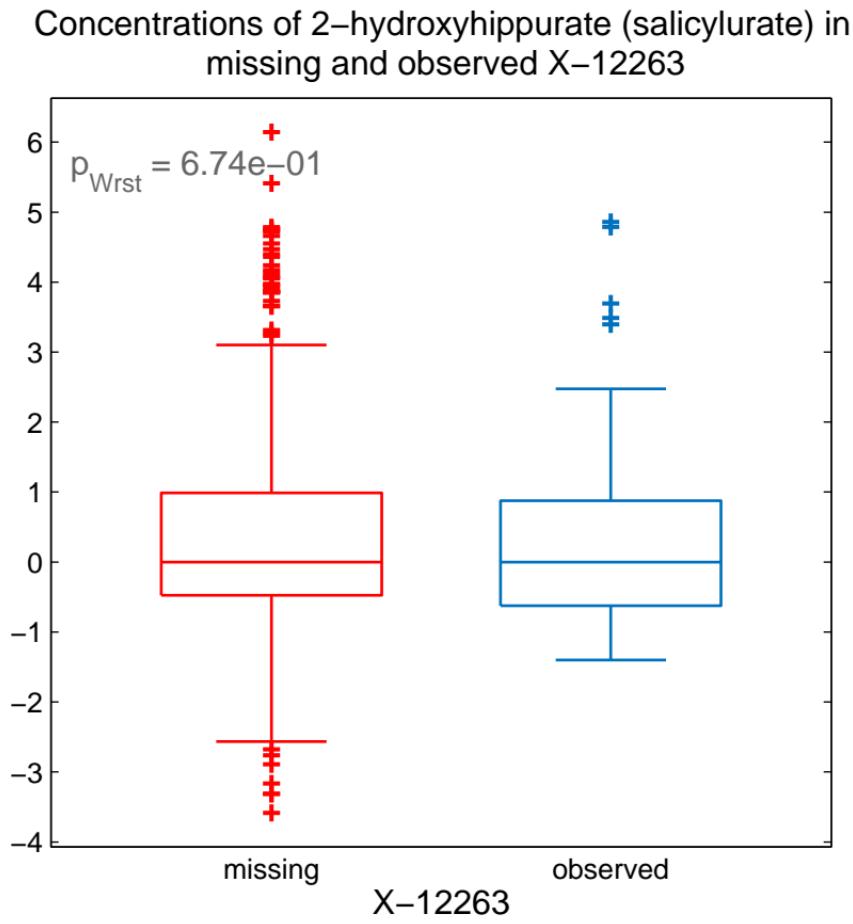
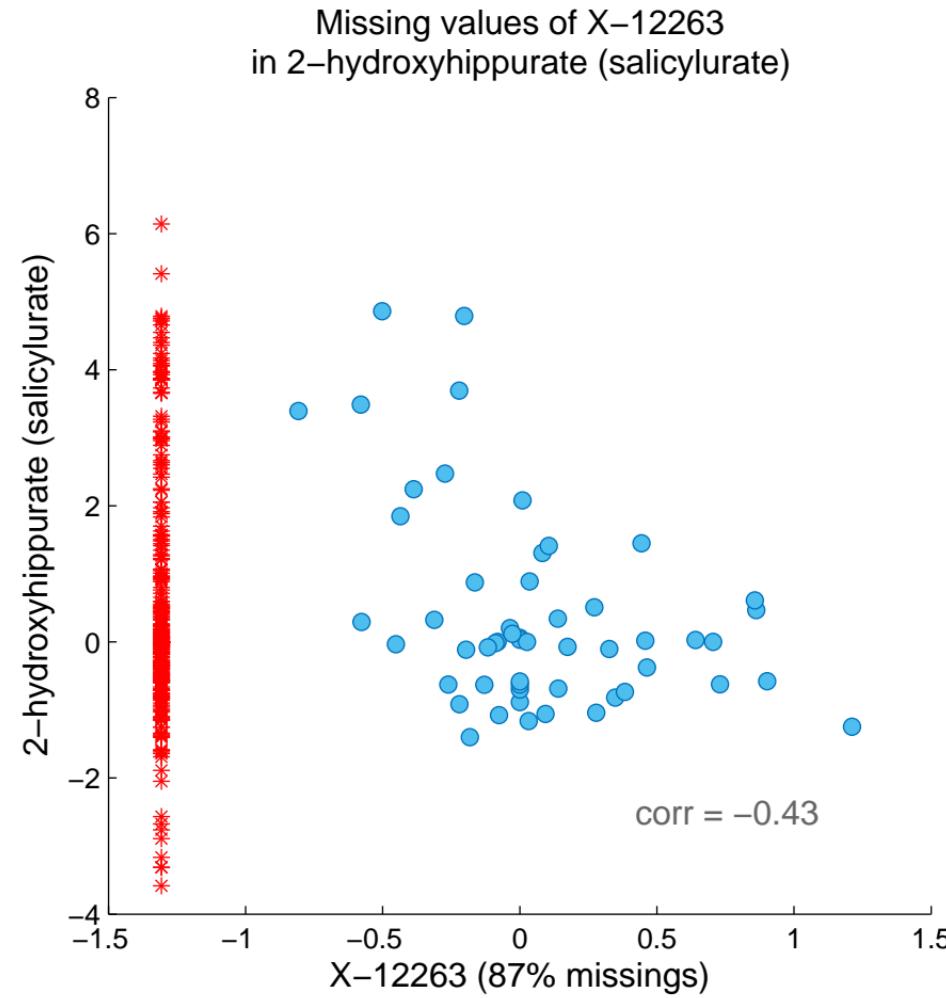


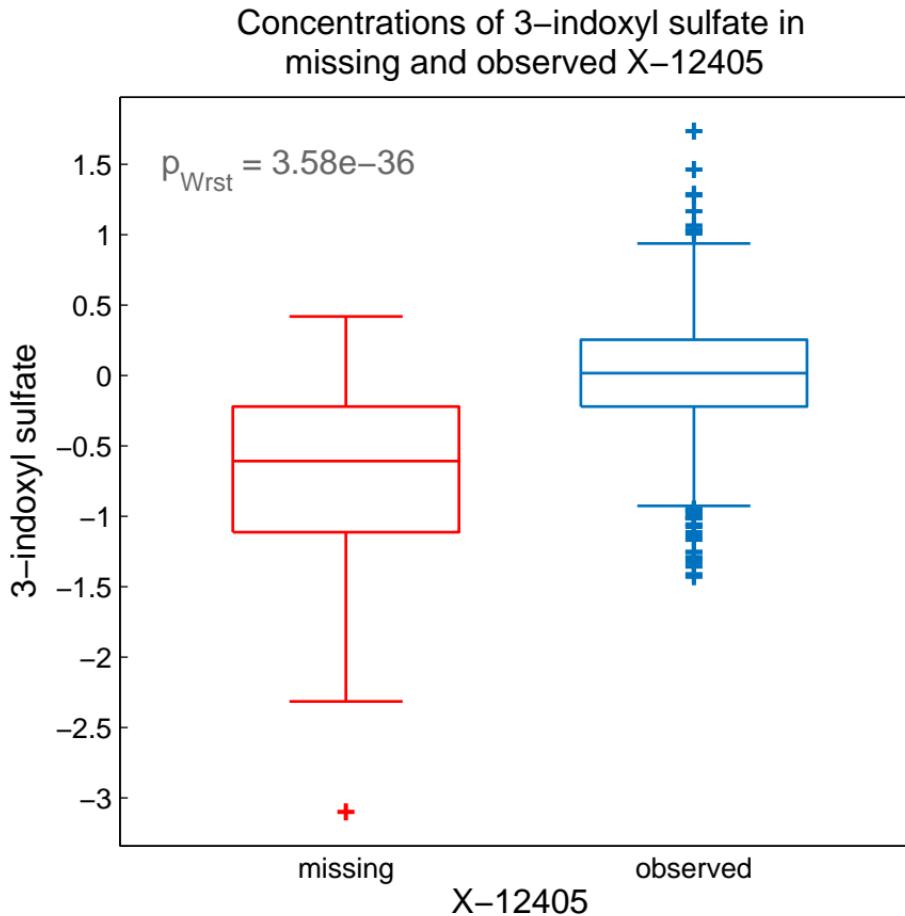
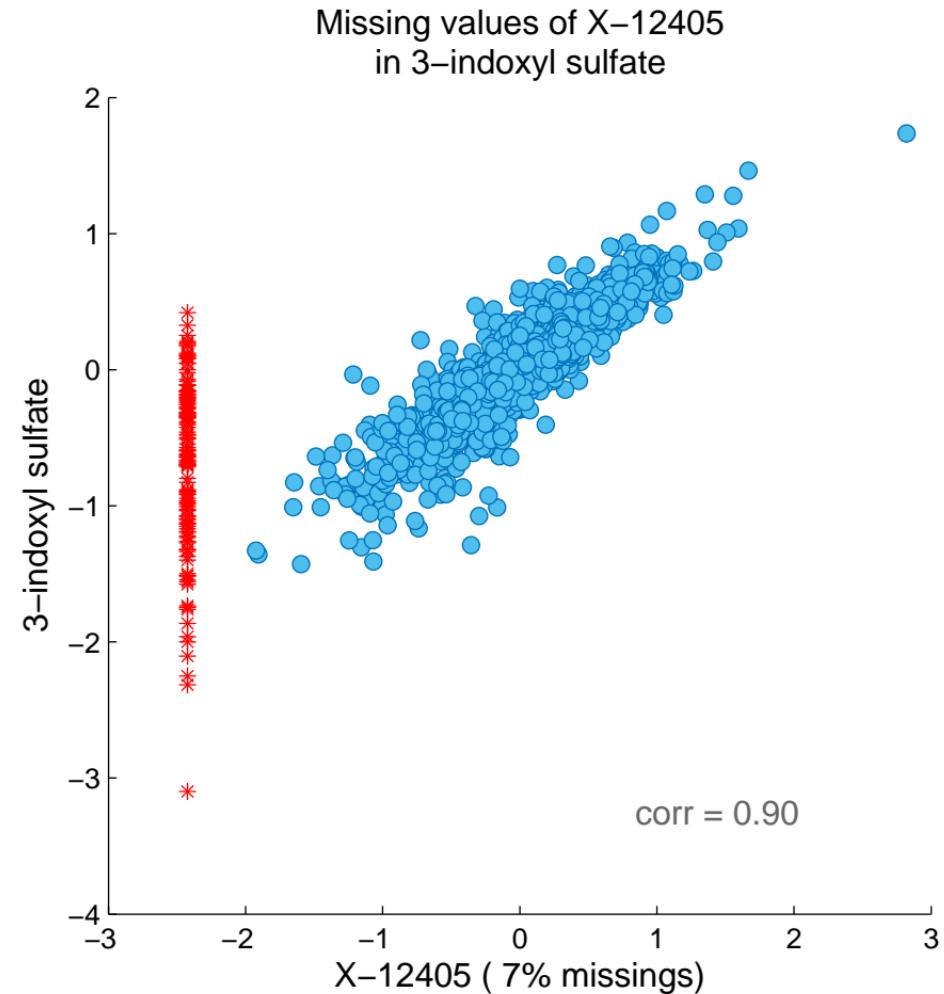


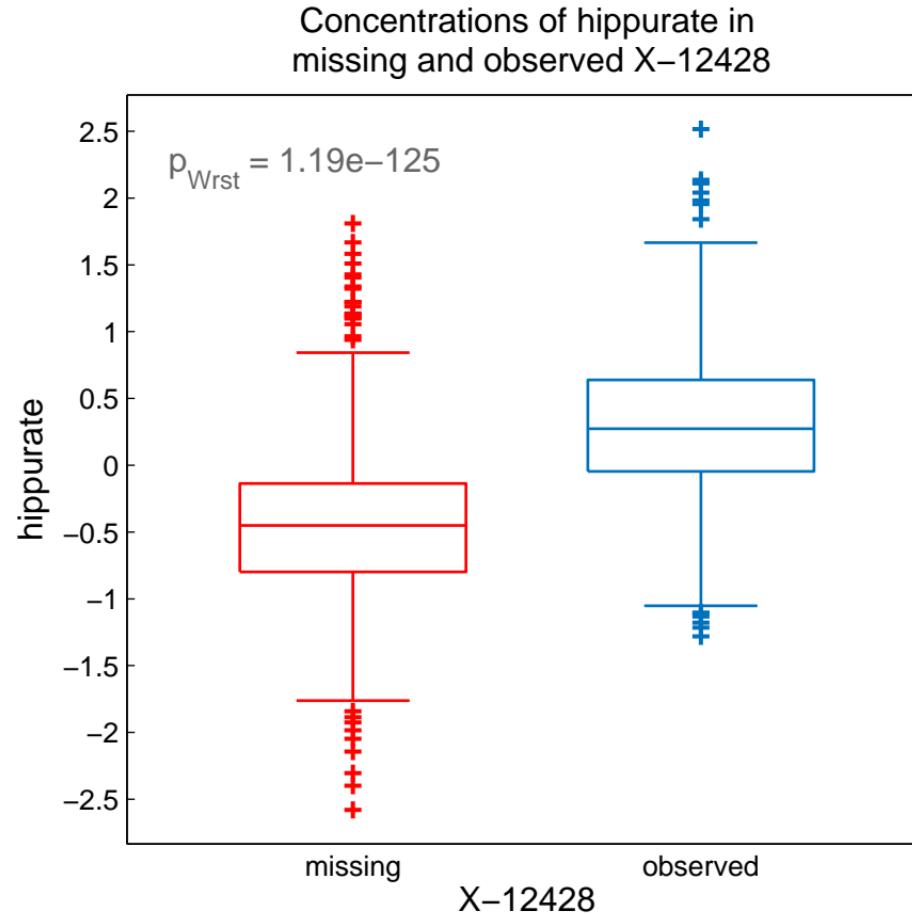
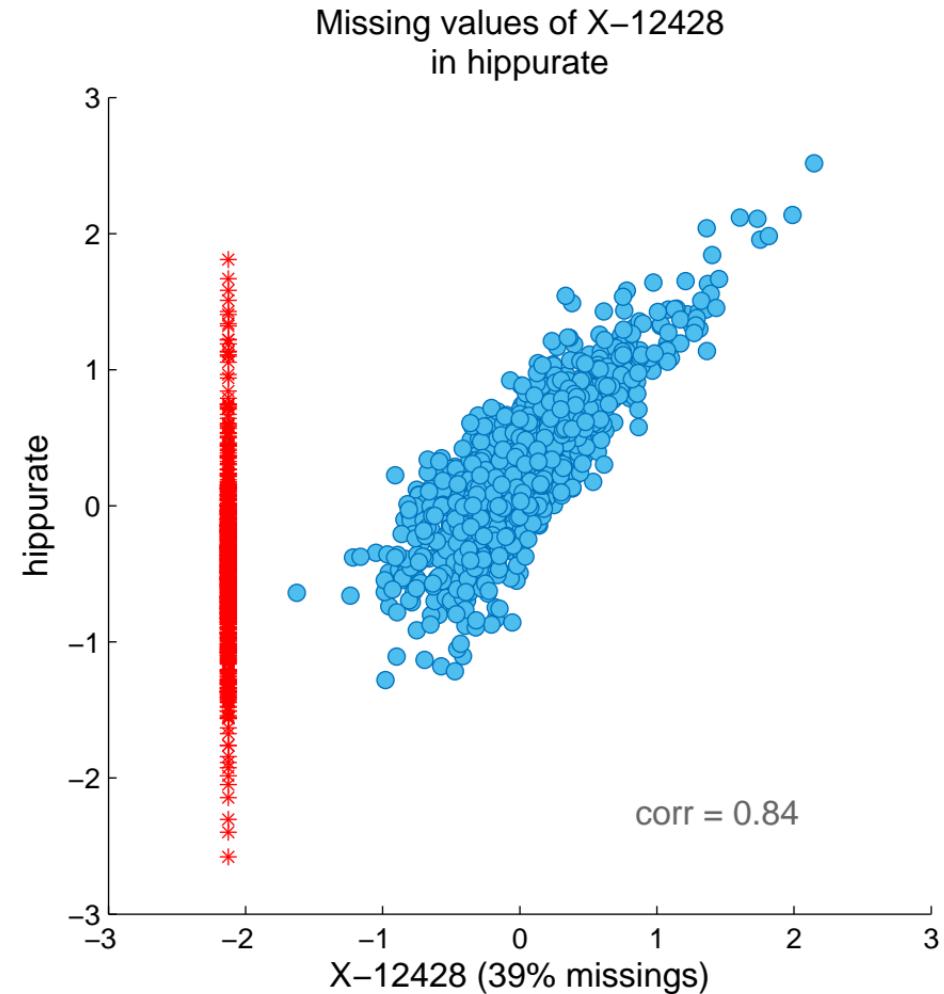


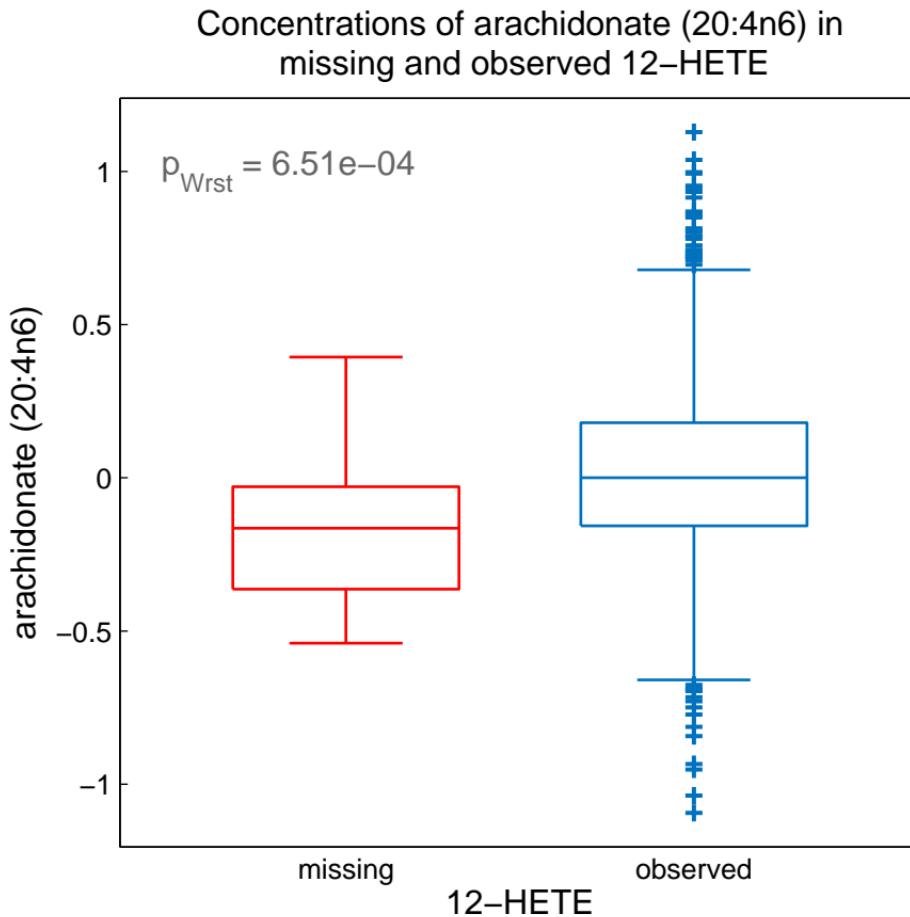
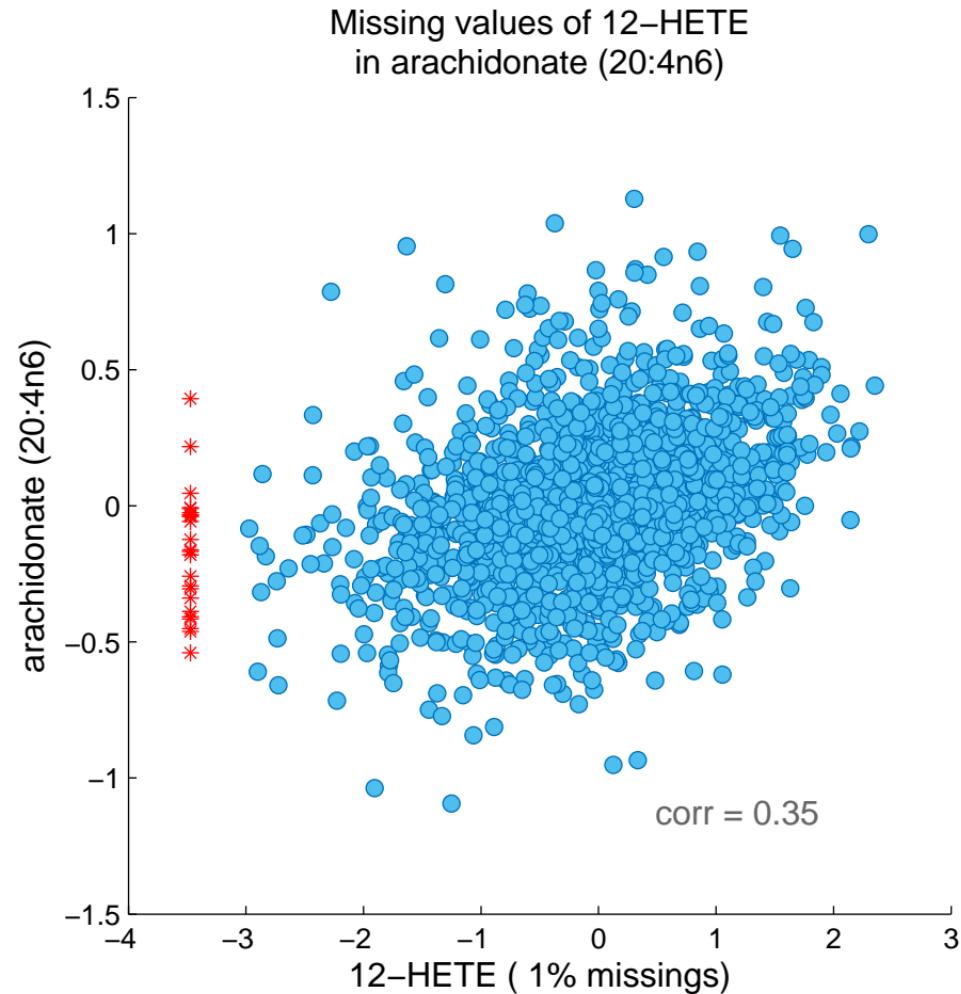




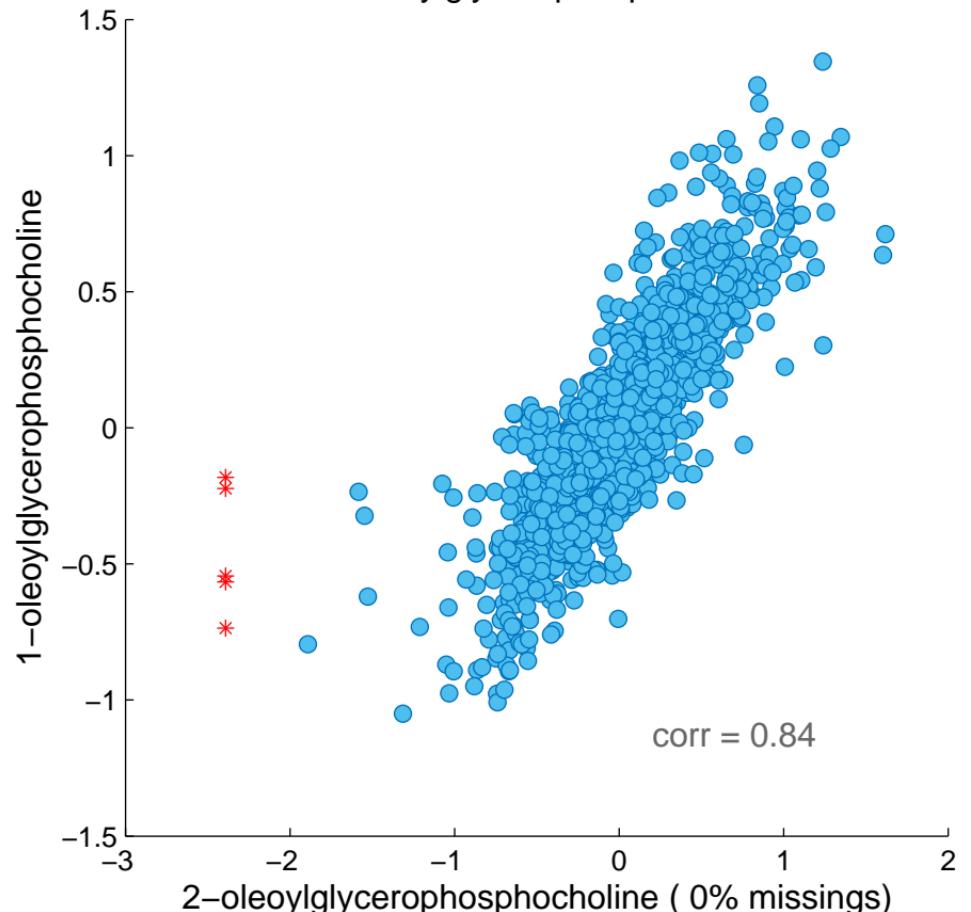




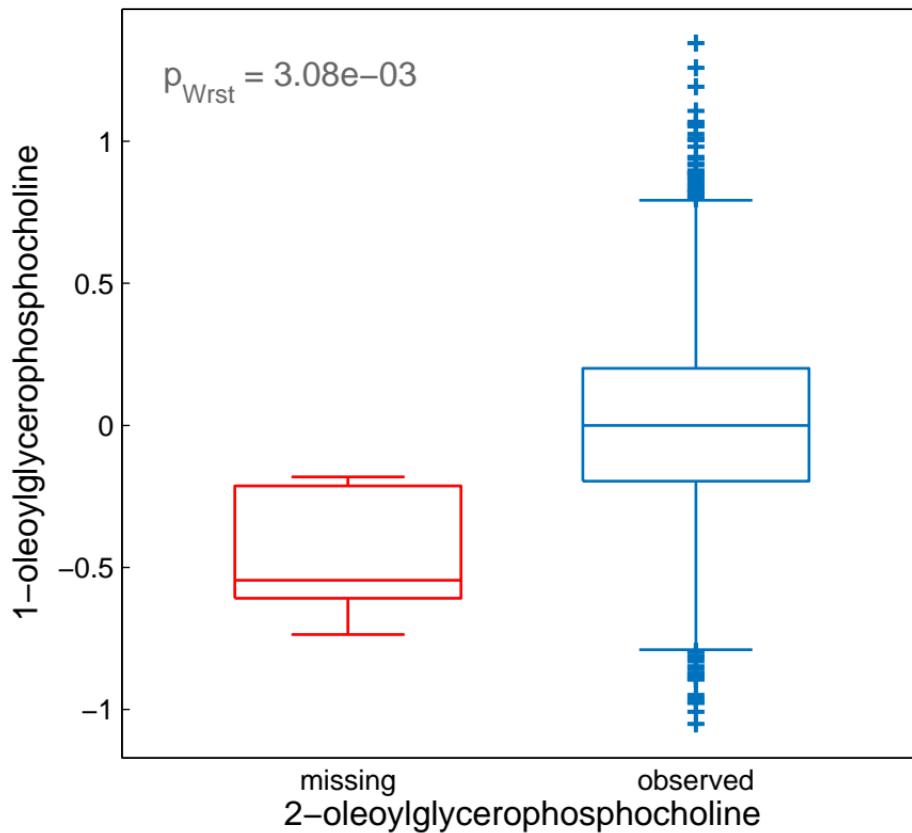


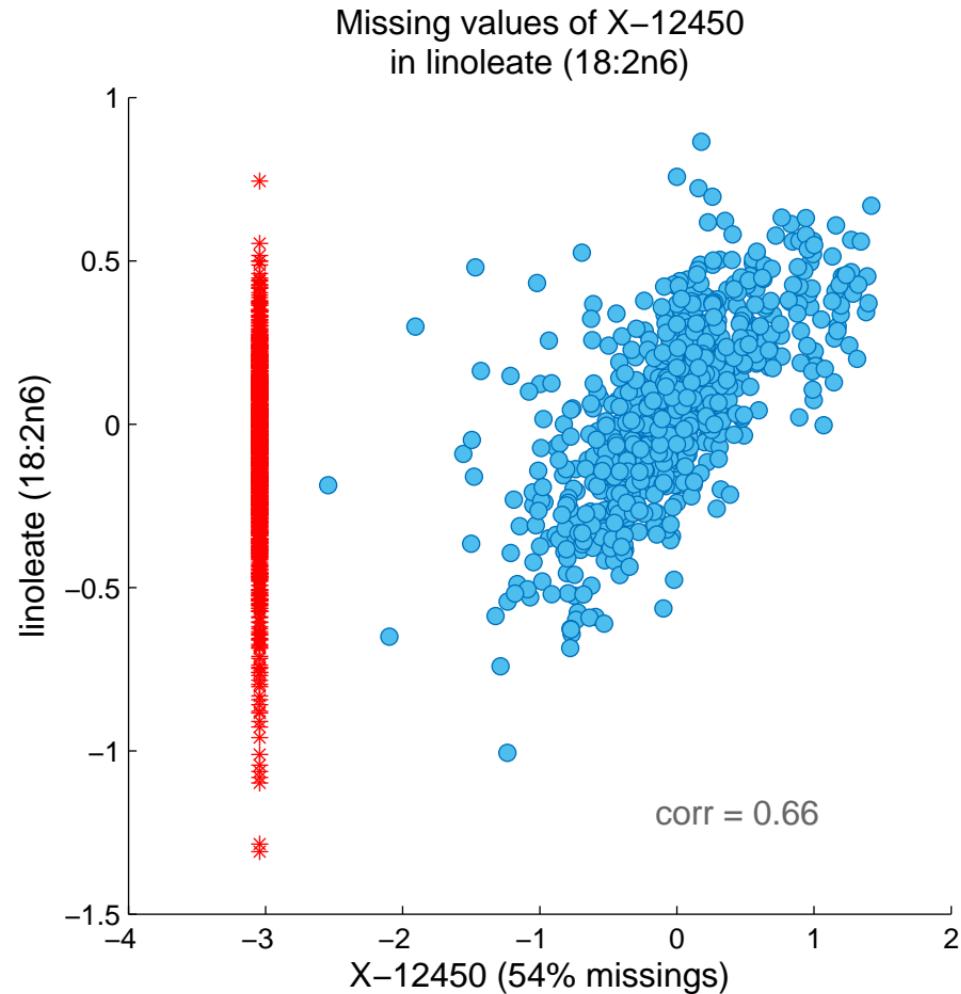


Missing values of 2-oleoylglycerophosphocholine
in 1-oleoylglycerophosphocholine

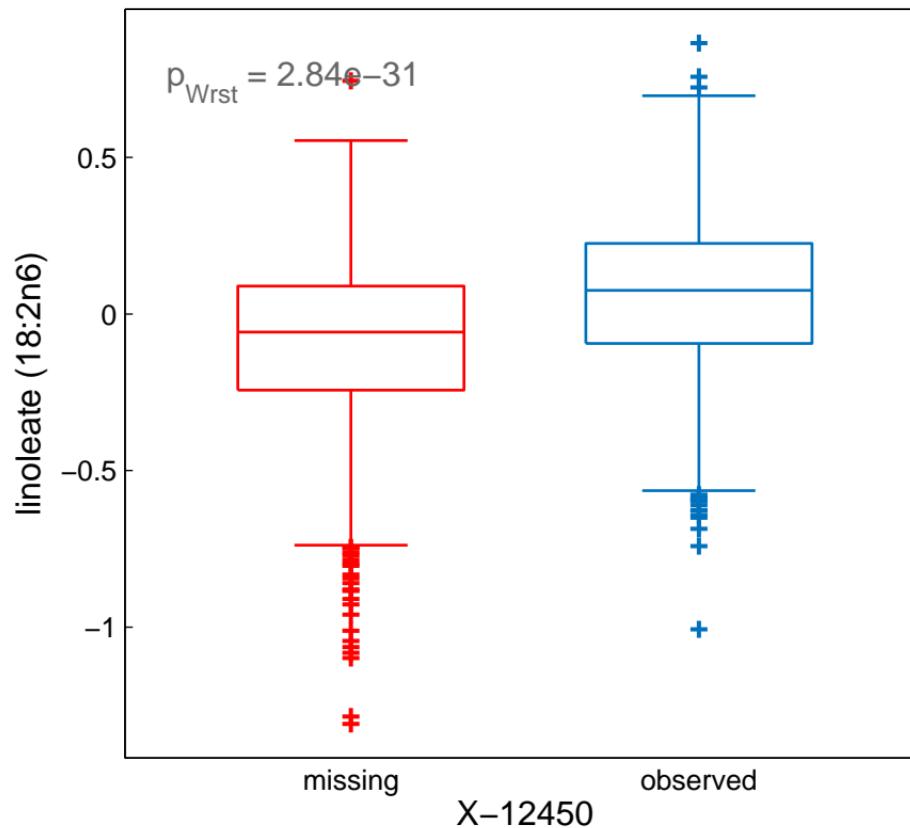


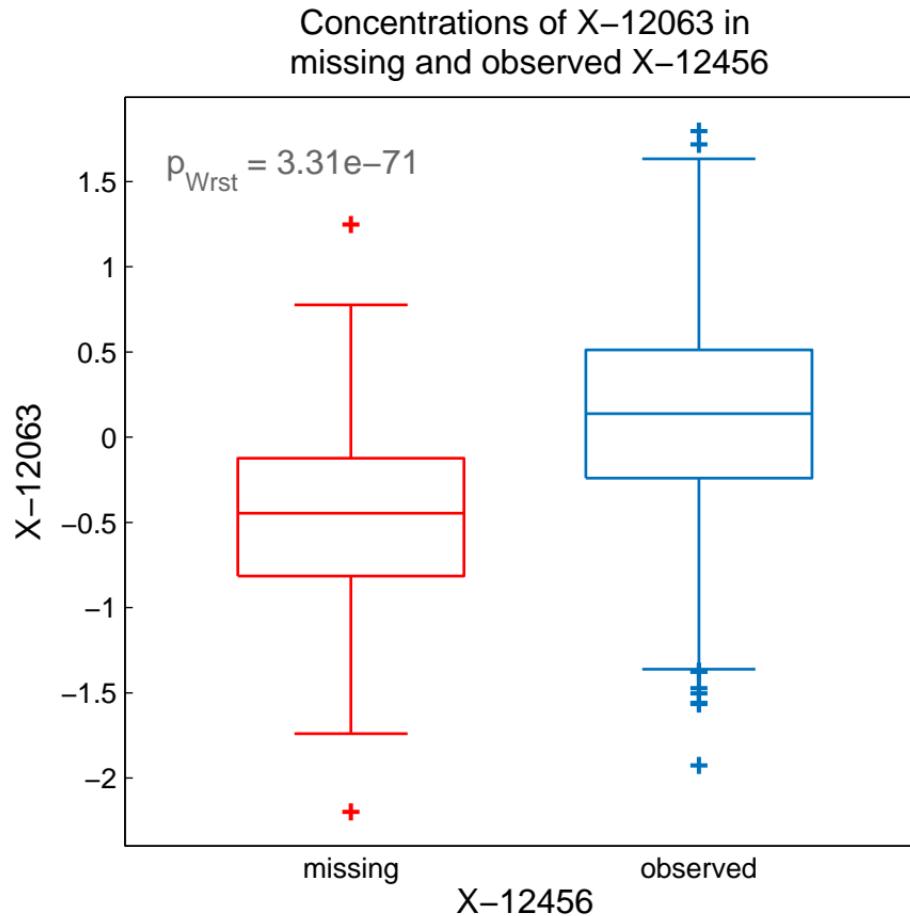
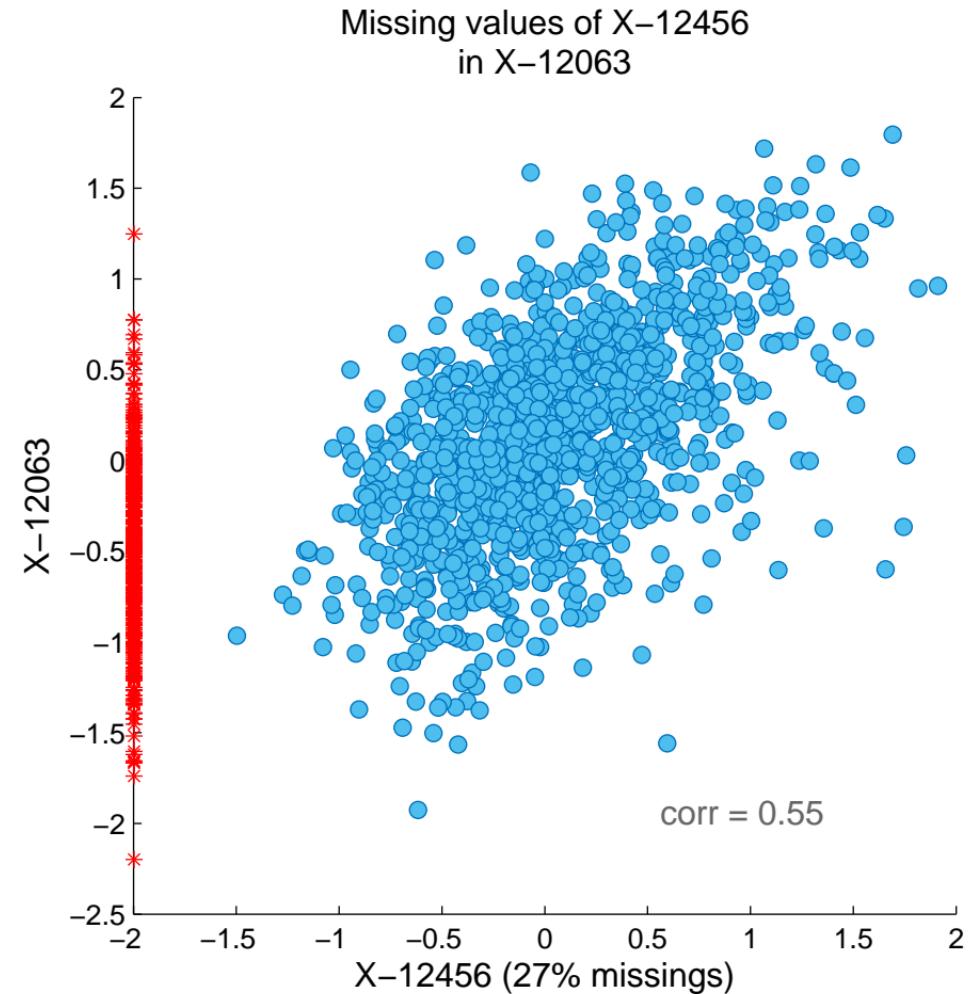
Concentrations of 1-oleoylglycerophosphocholine in
missing and observed 2-oleoylglycerophosphocholine

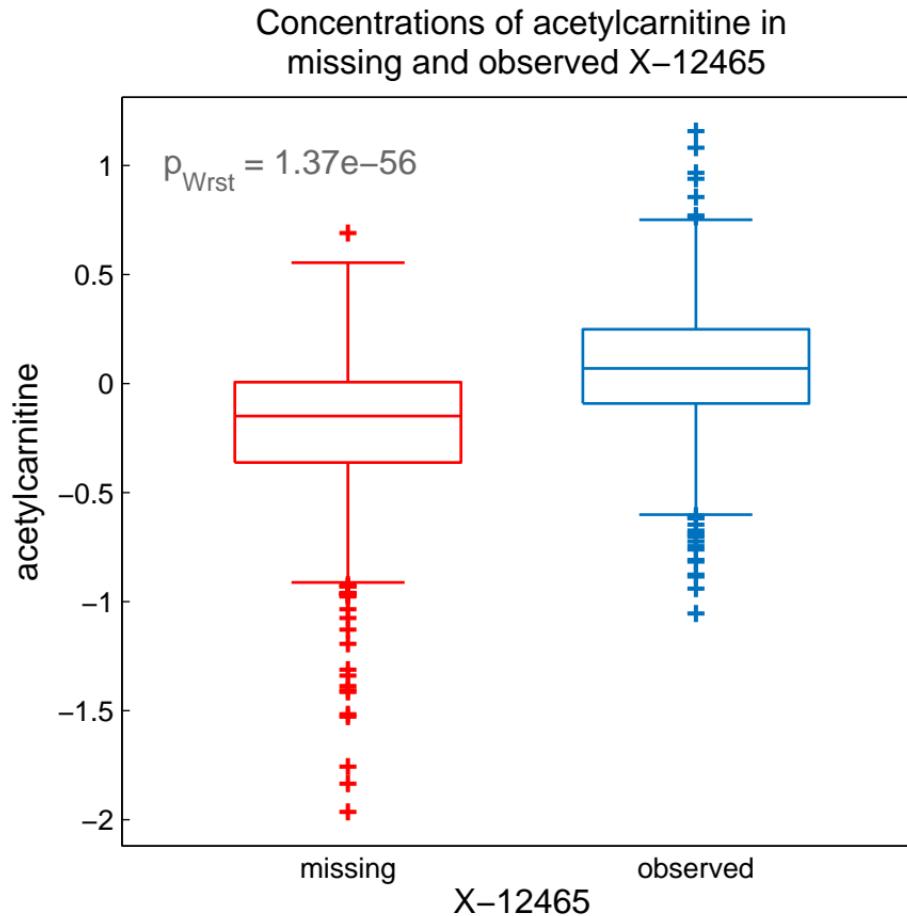
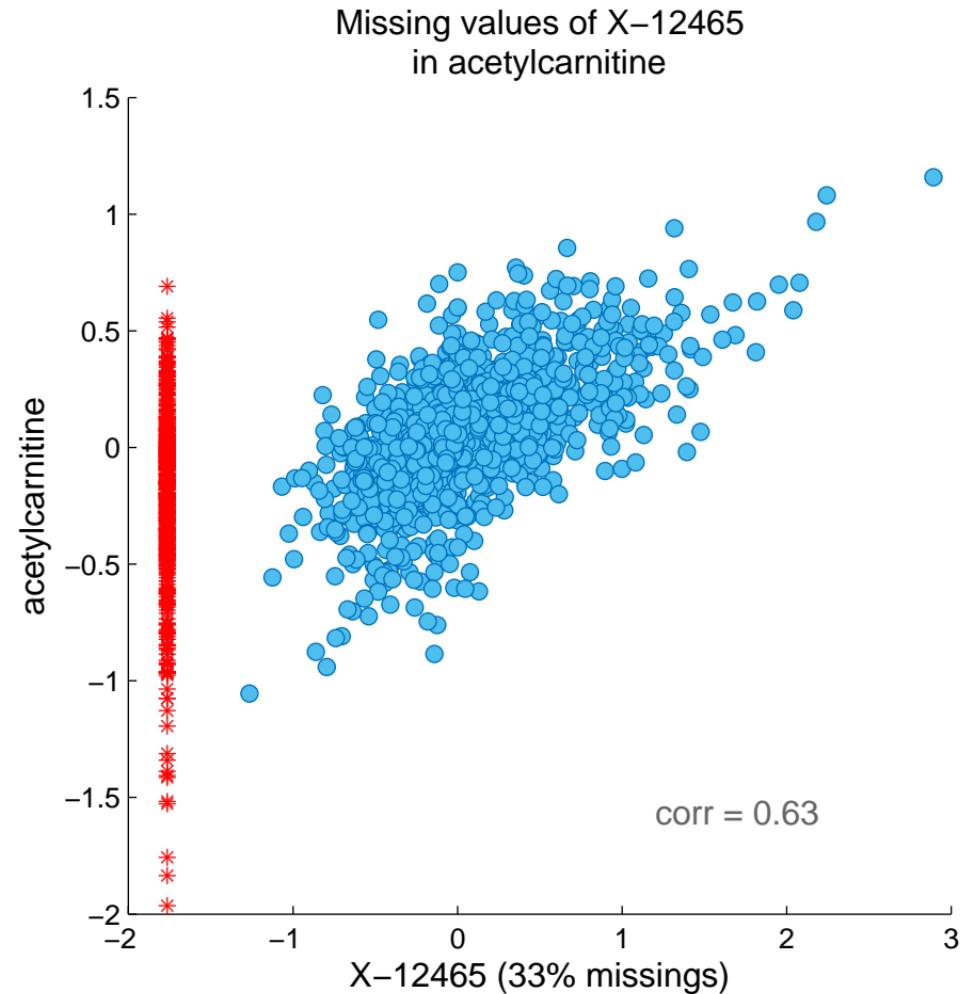


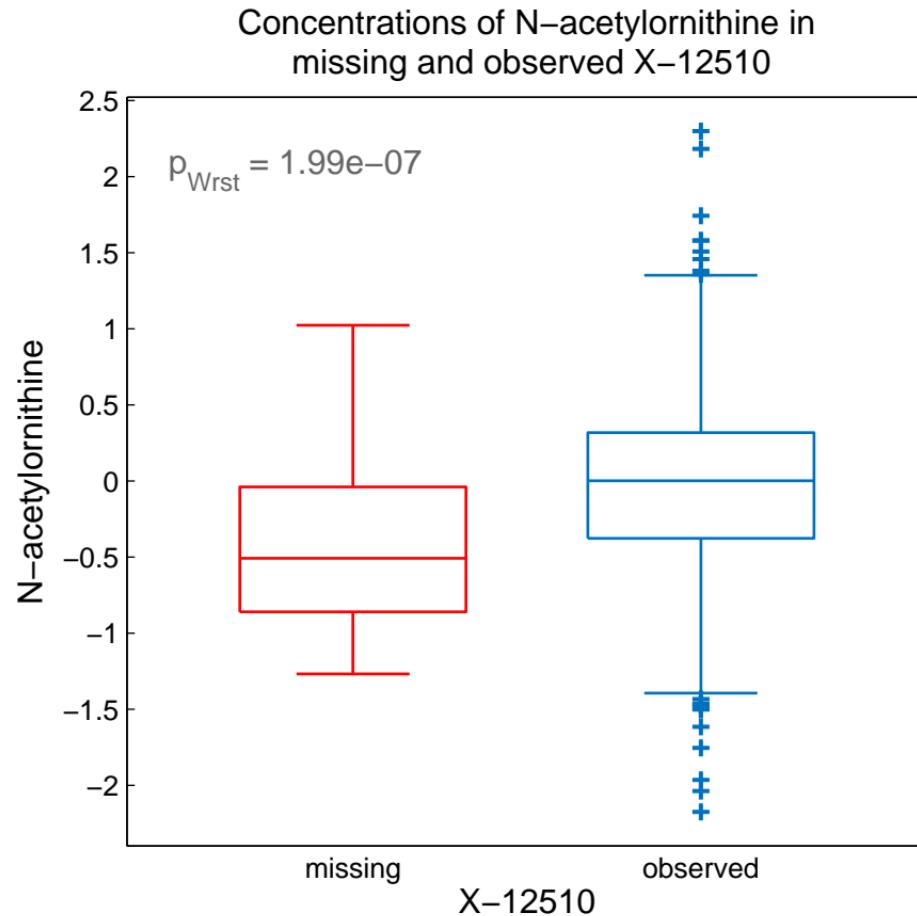
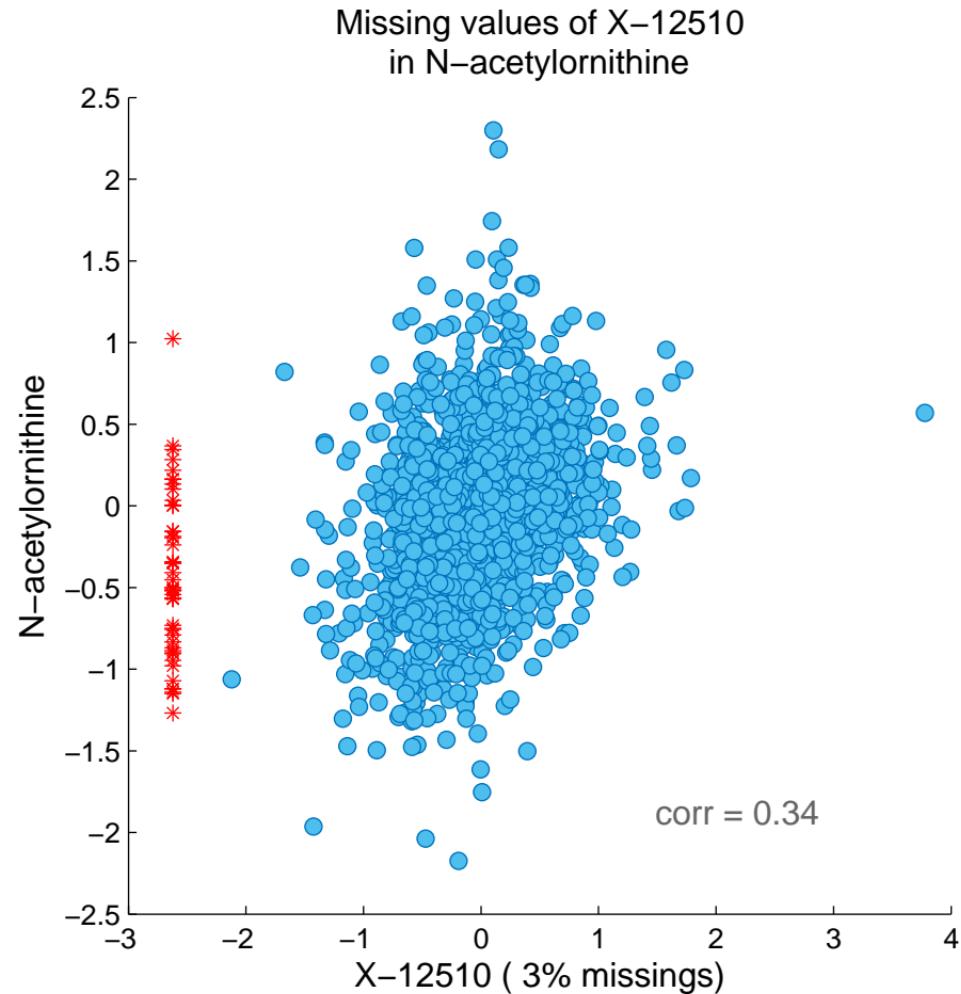


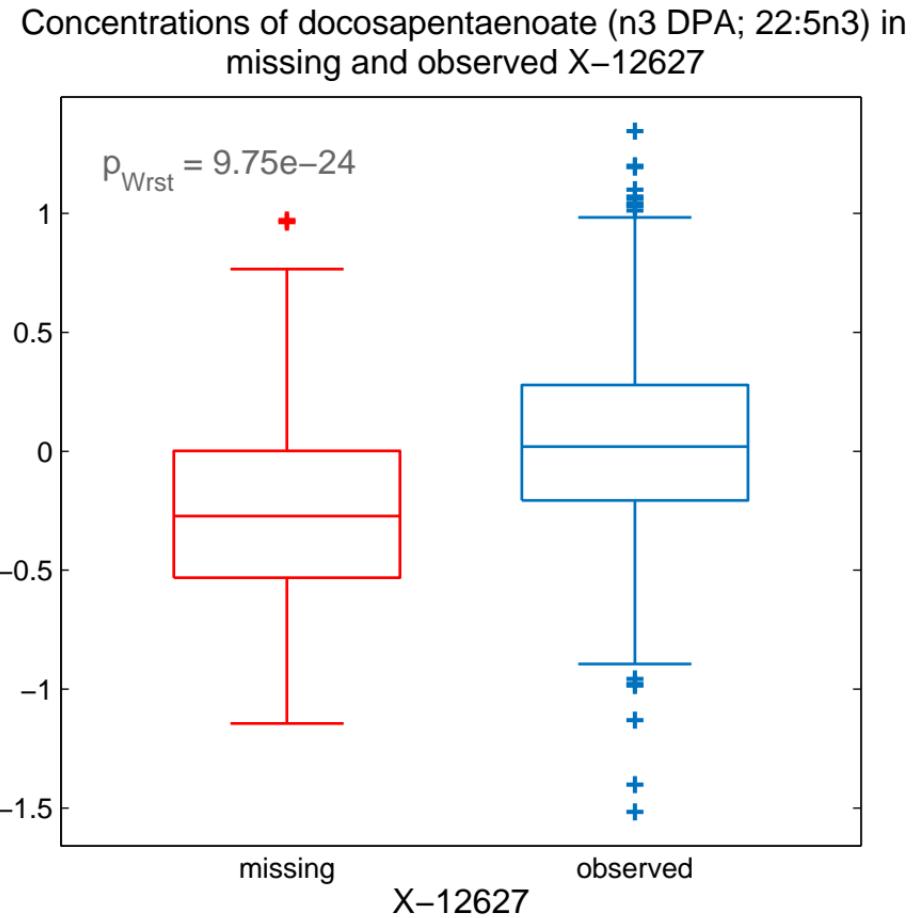
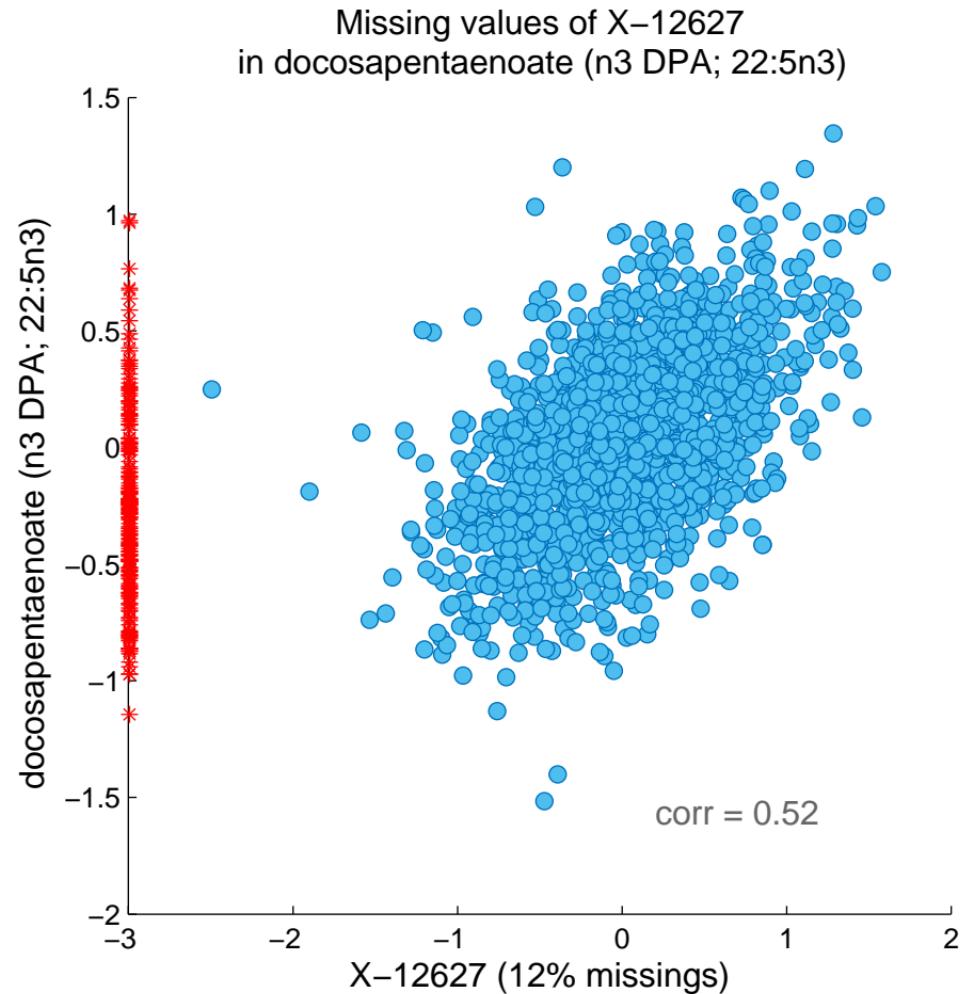
Concentrations of linoleate (18:2n6) in
missing and observed X-12450

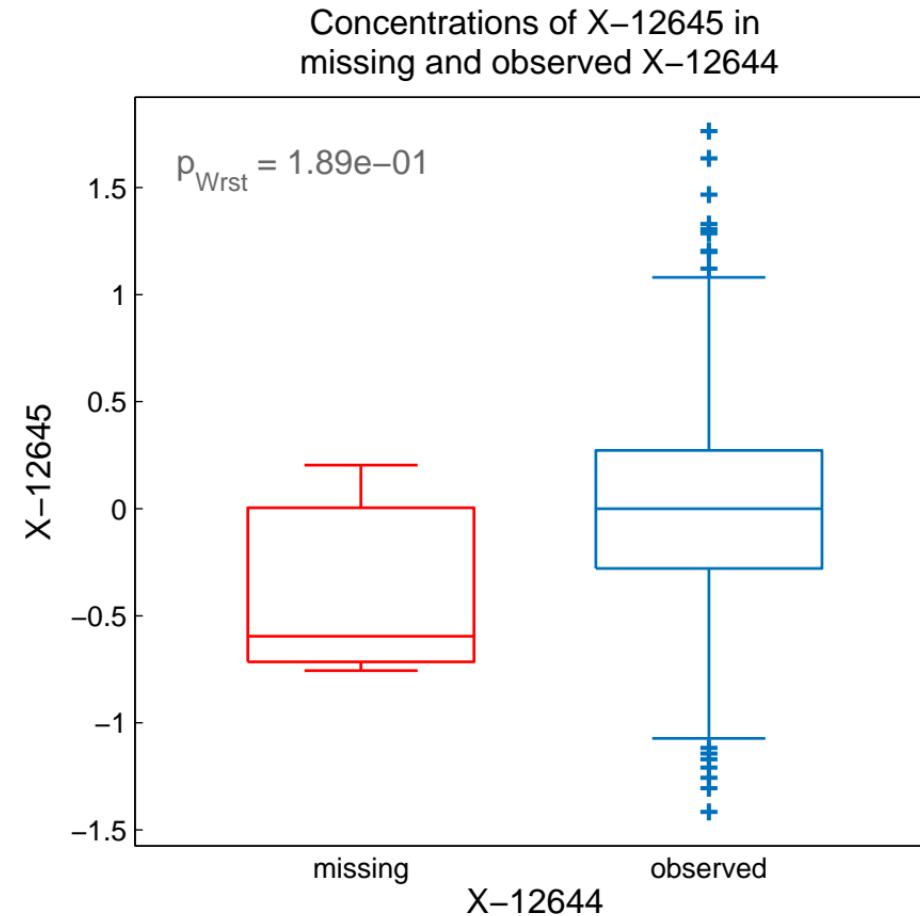
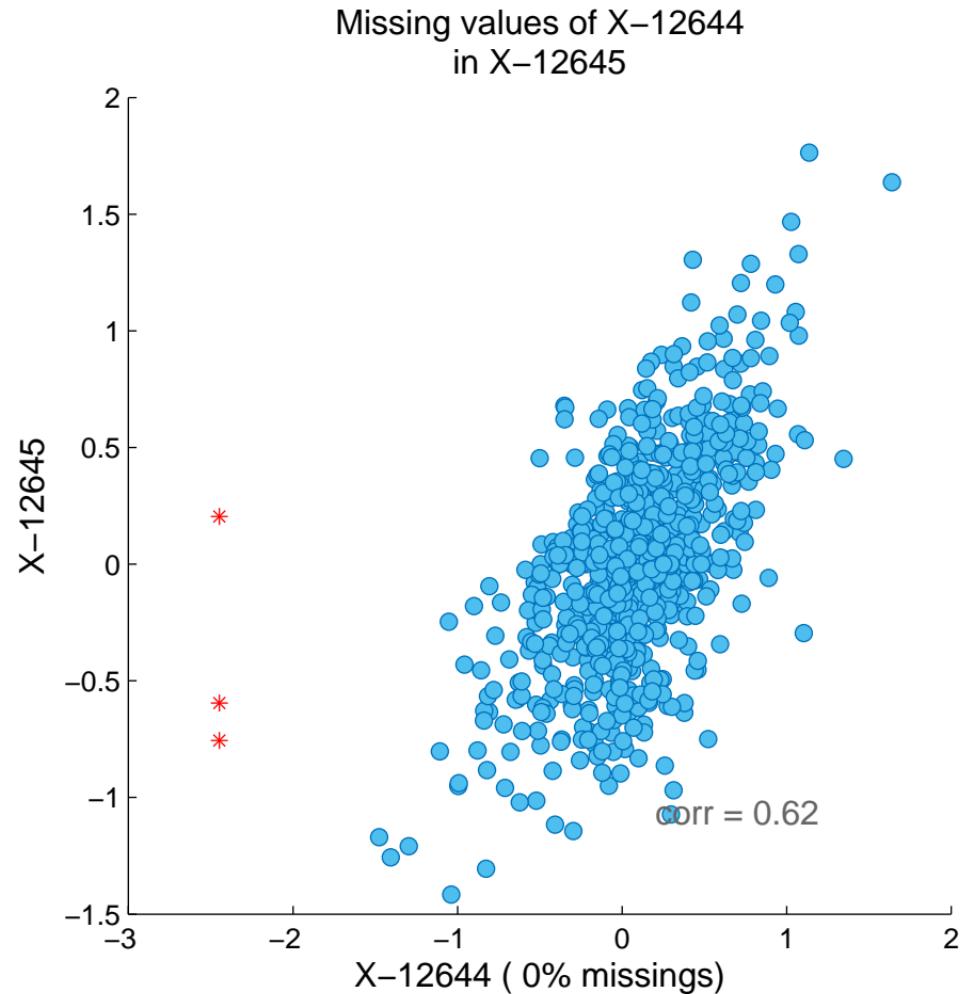


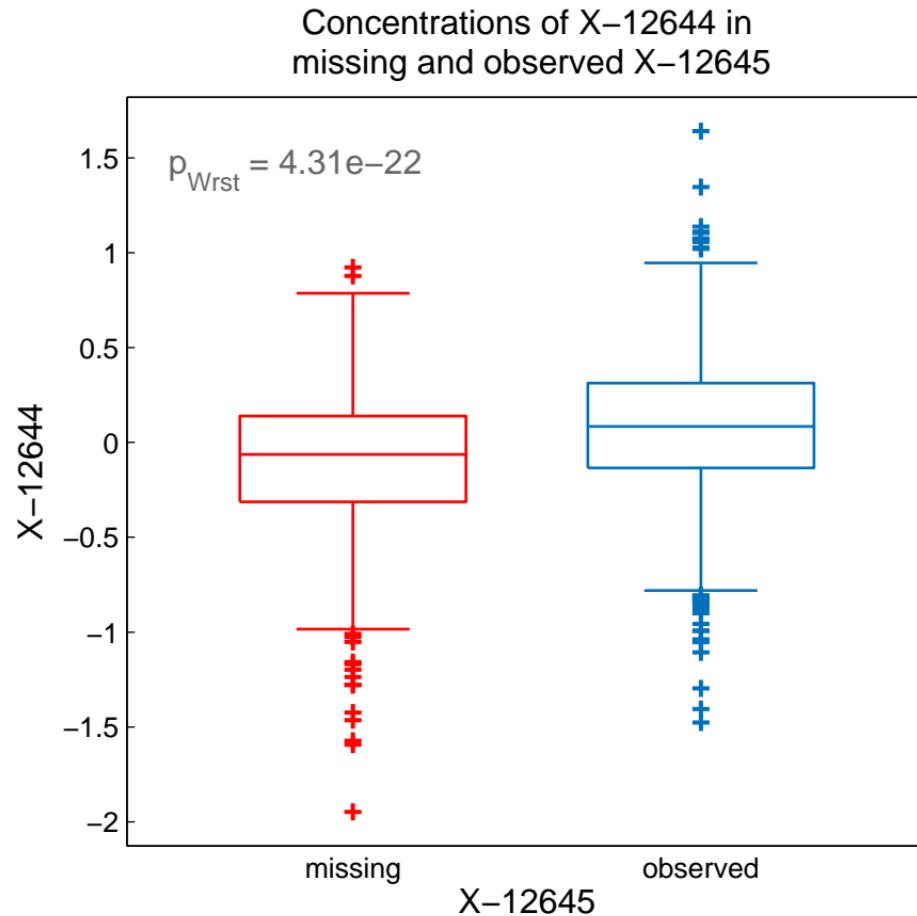
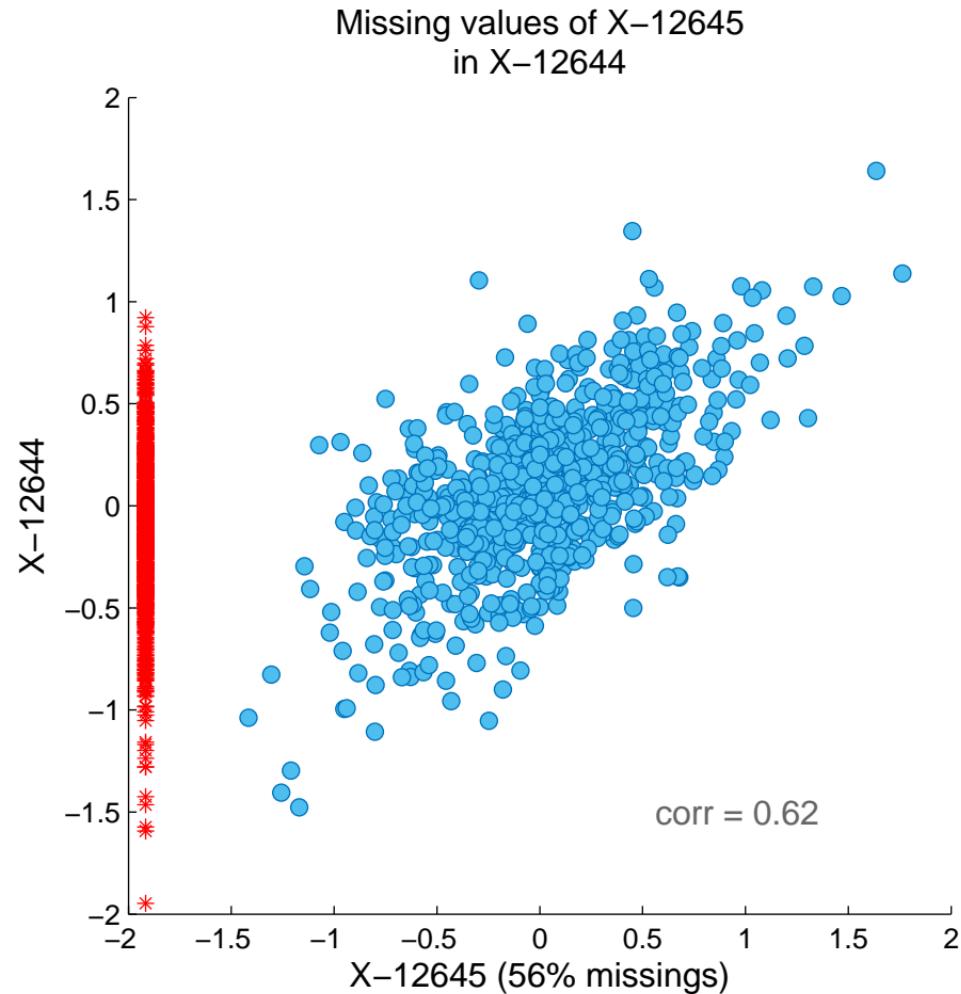


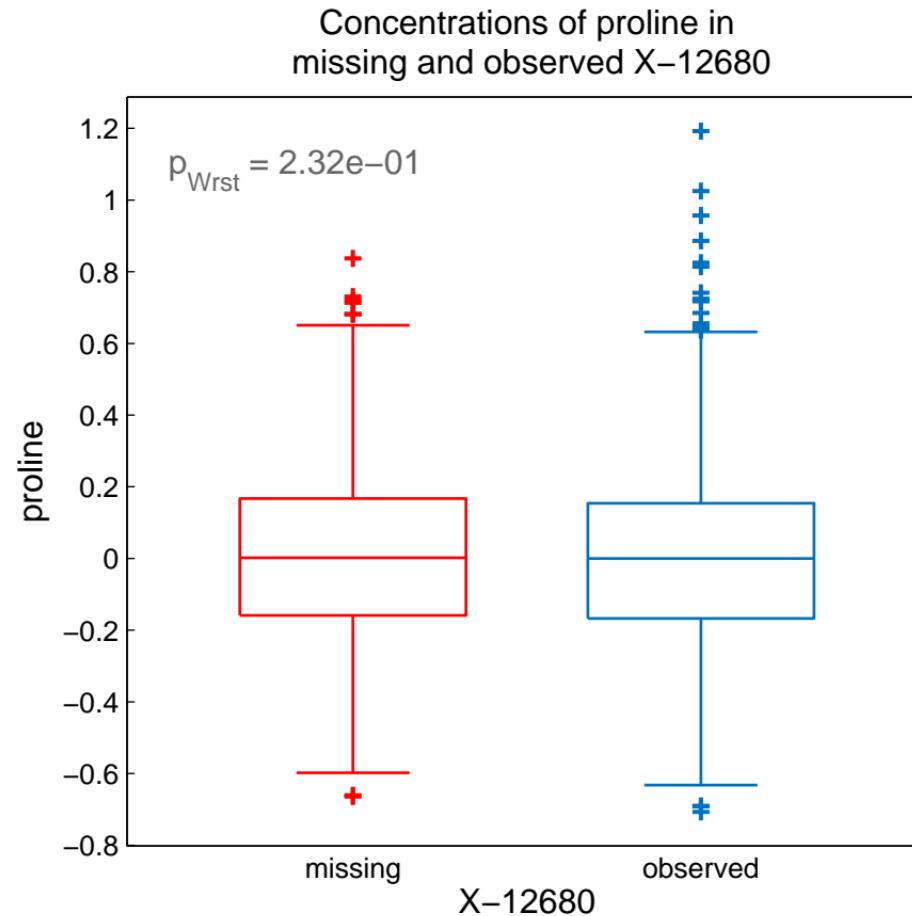
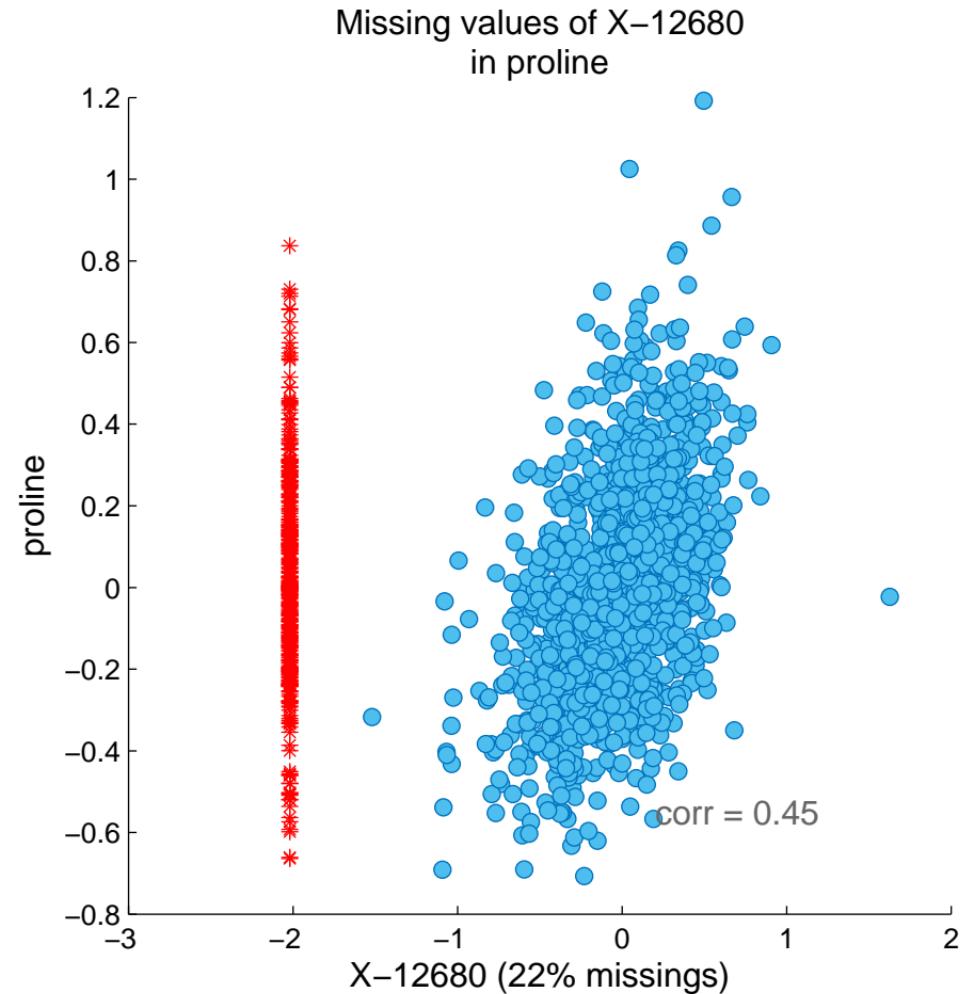


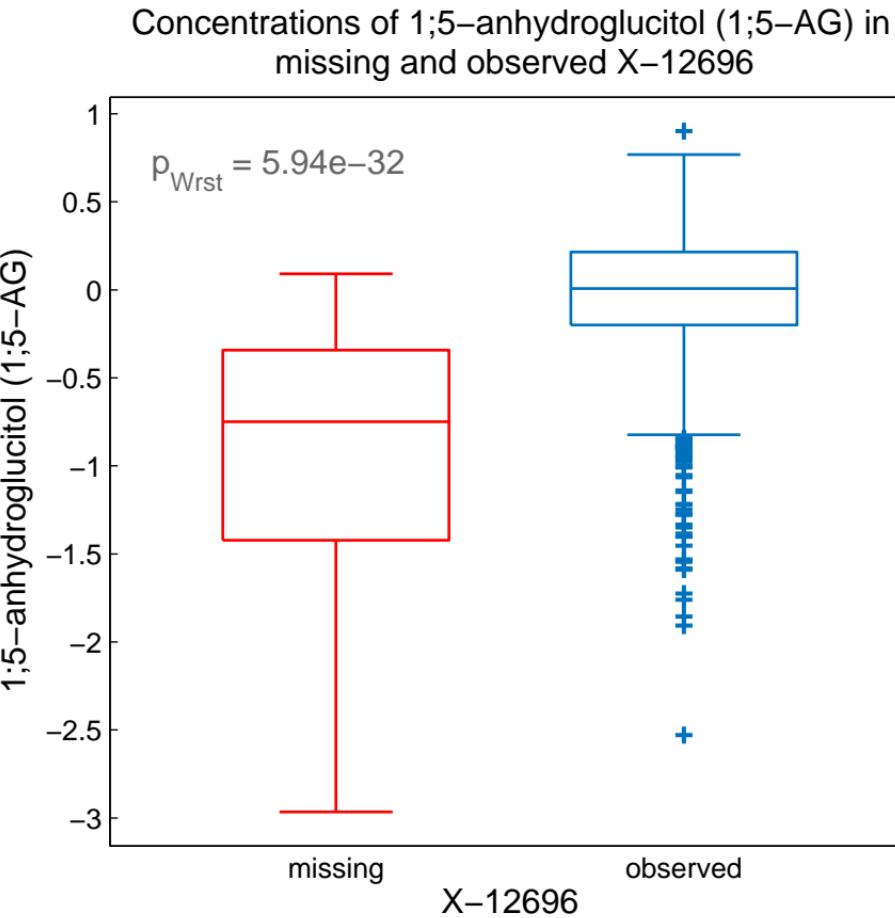
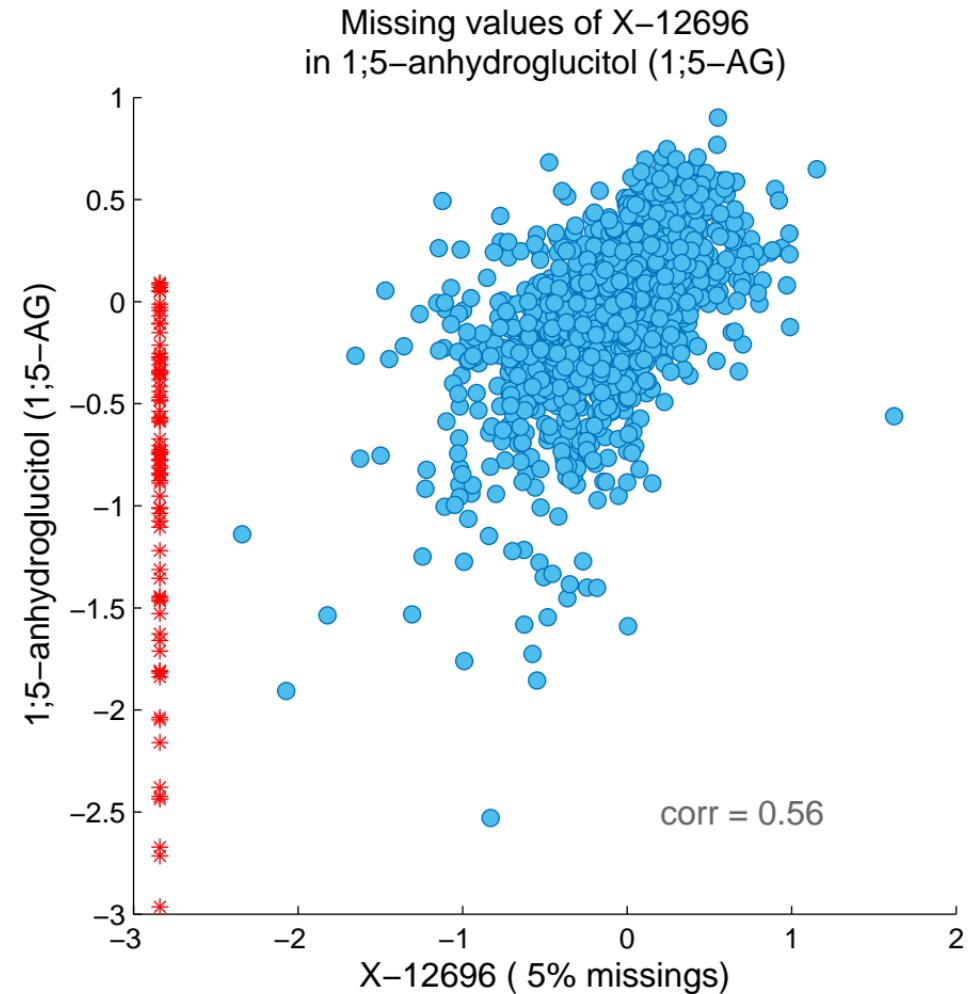


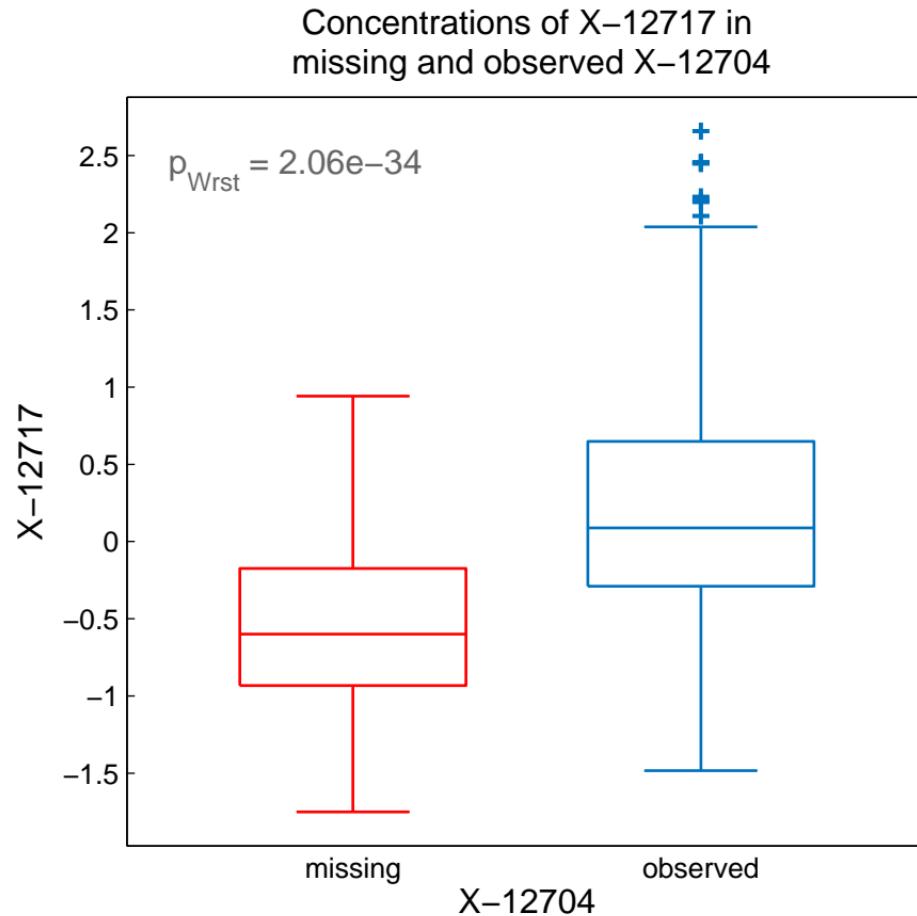
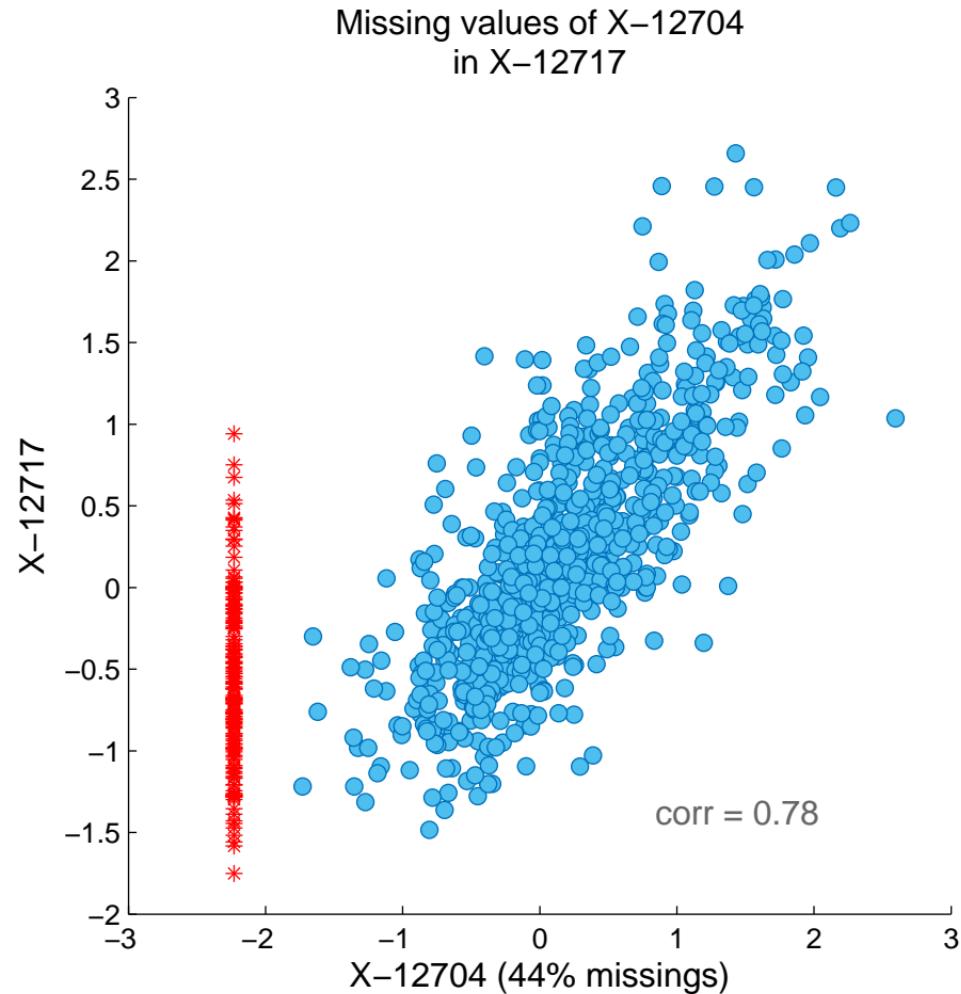


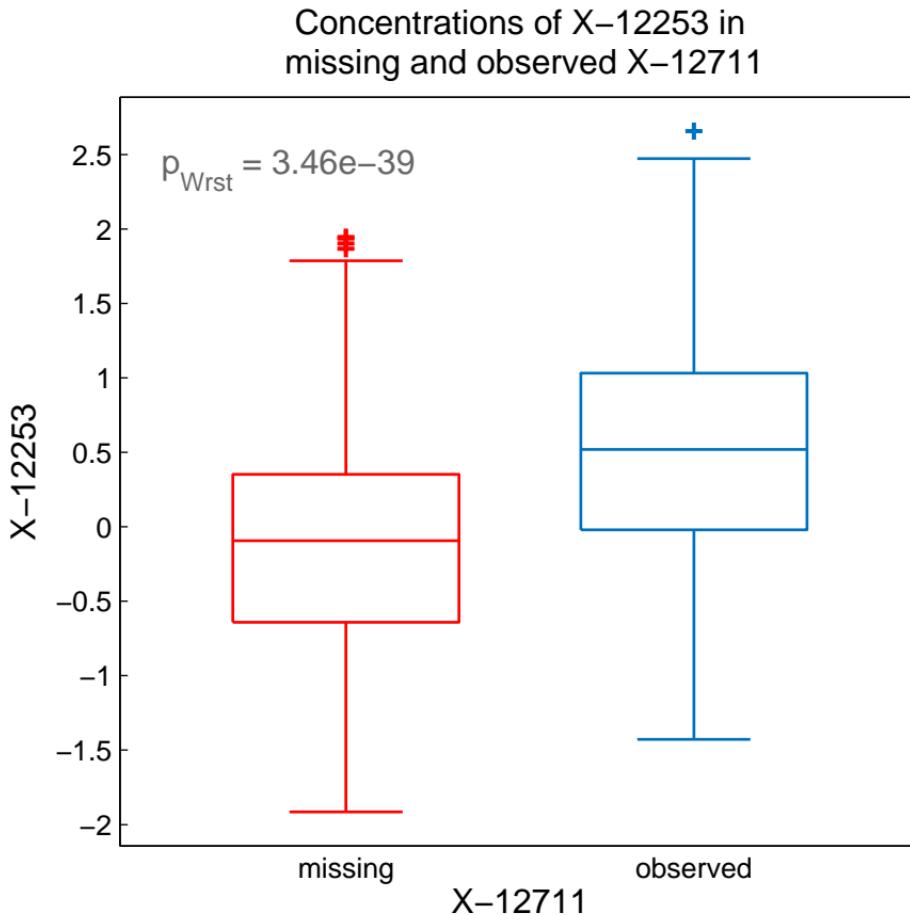
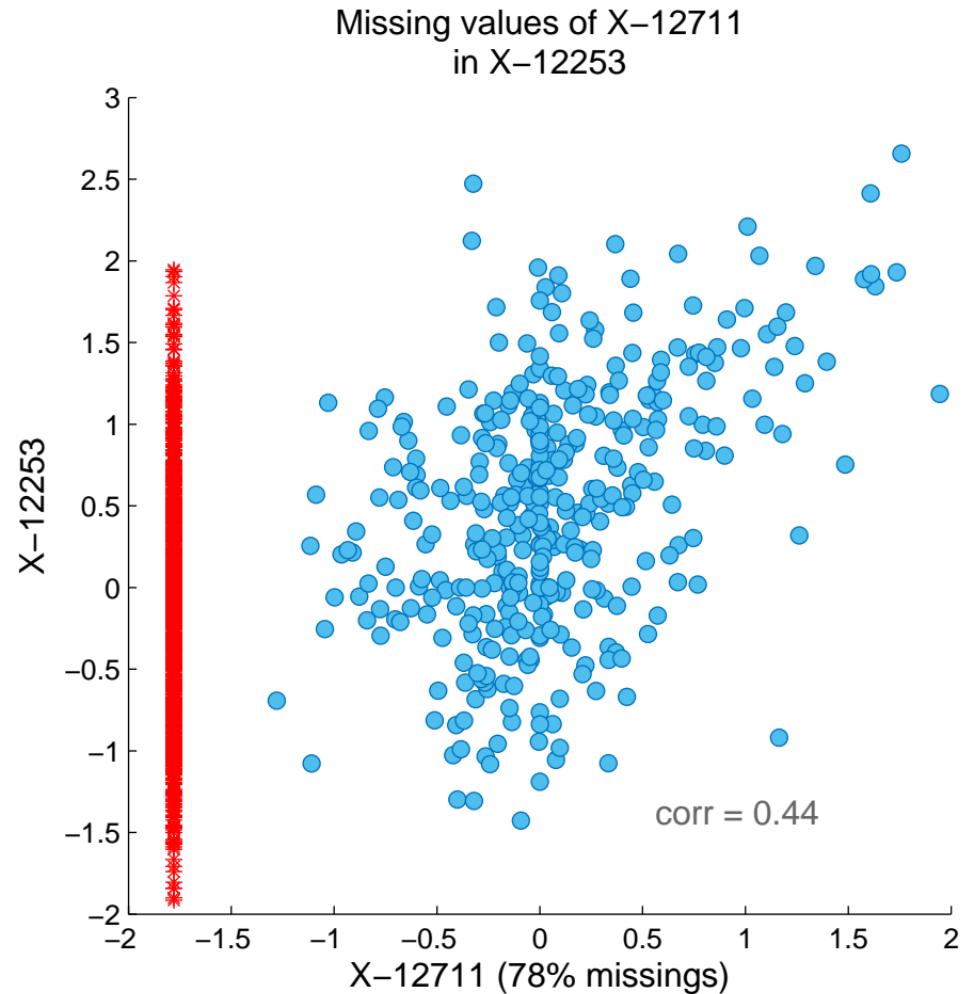


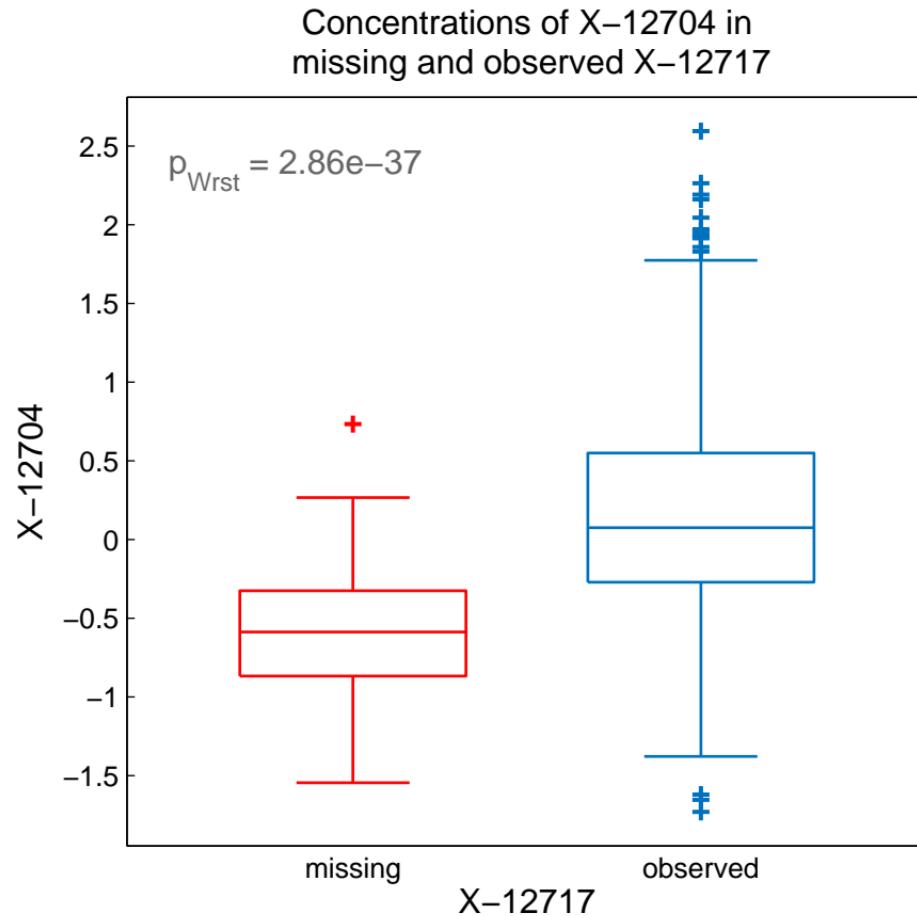
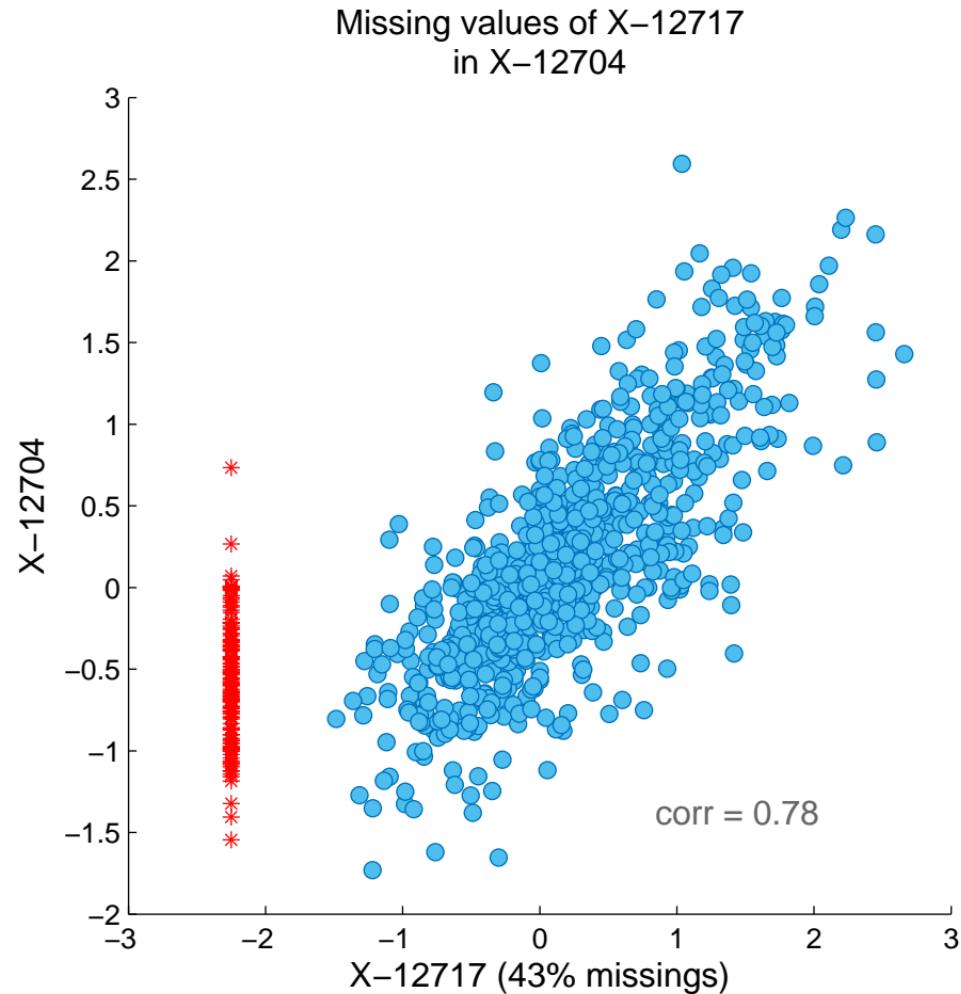


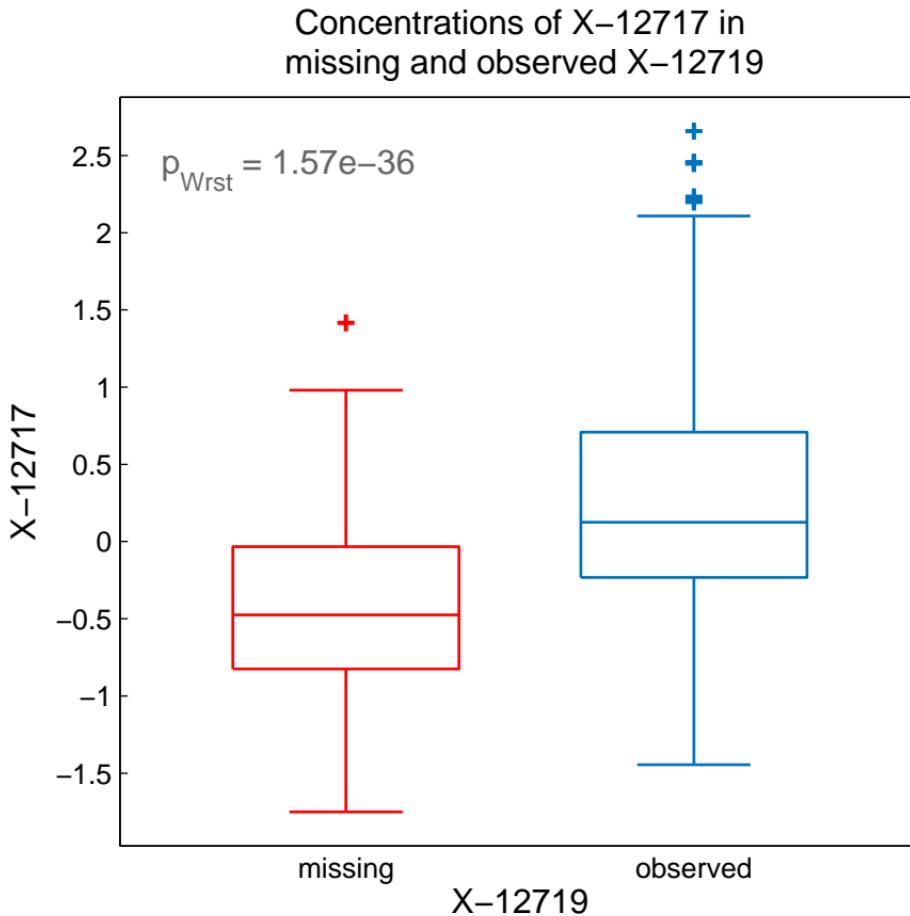
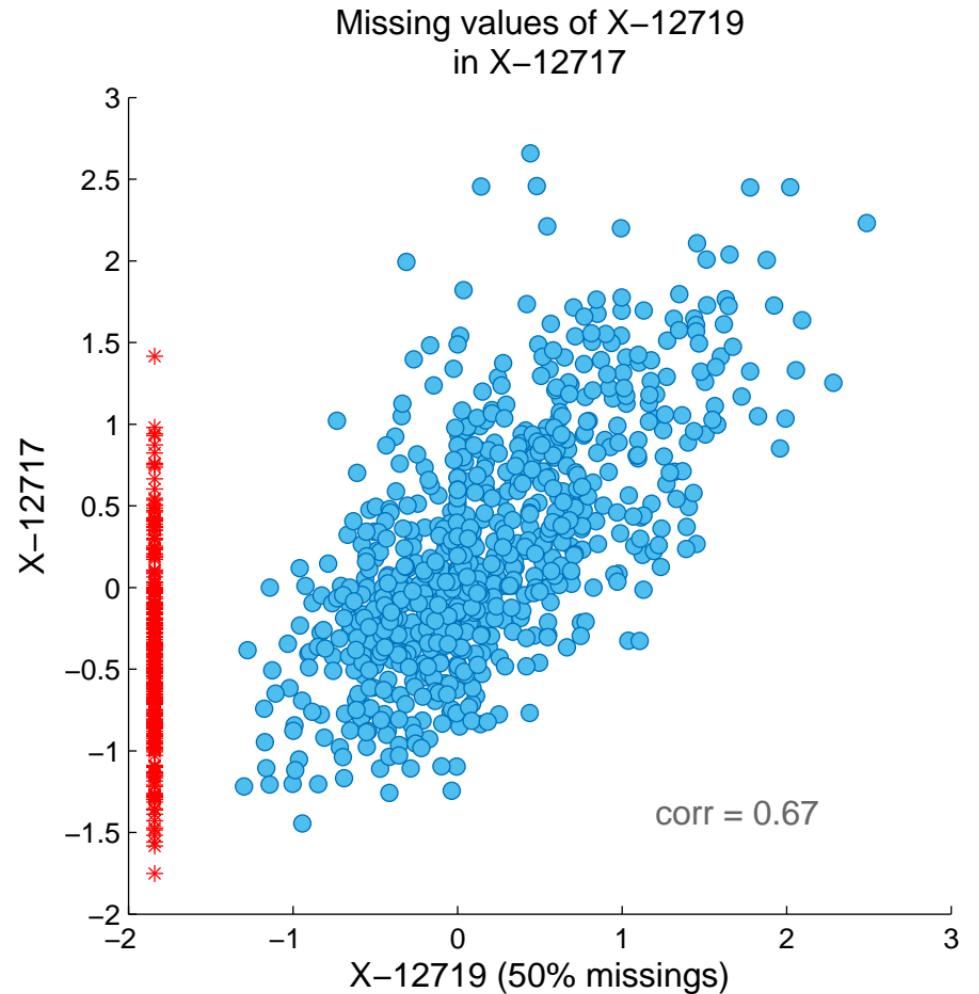


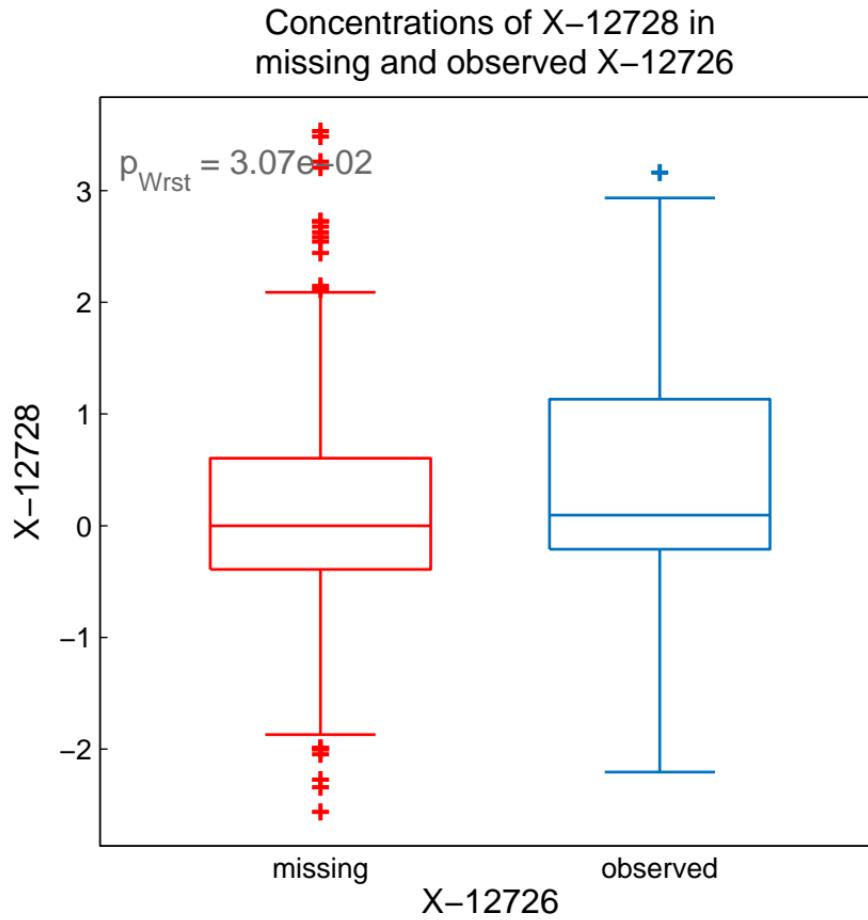
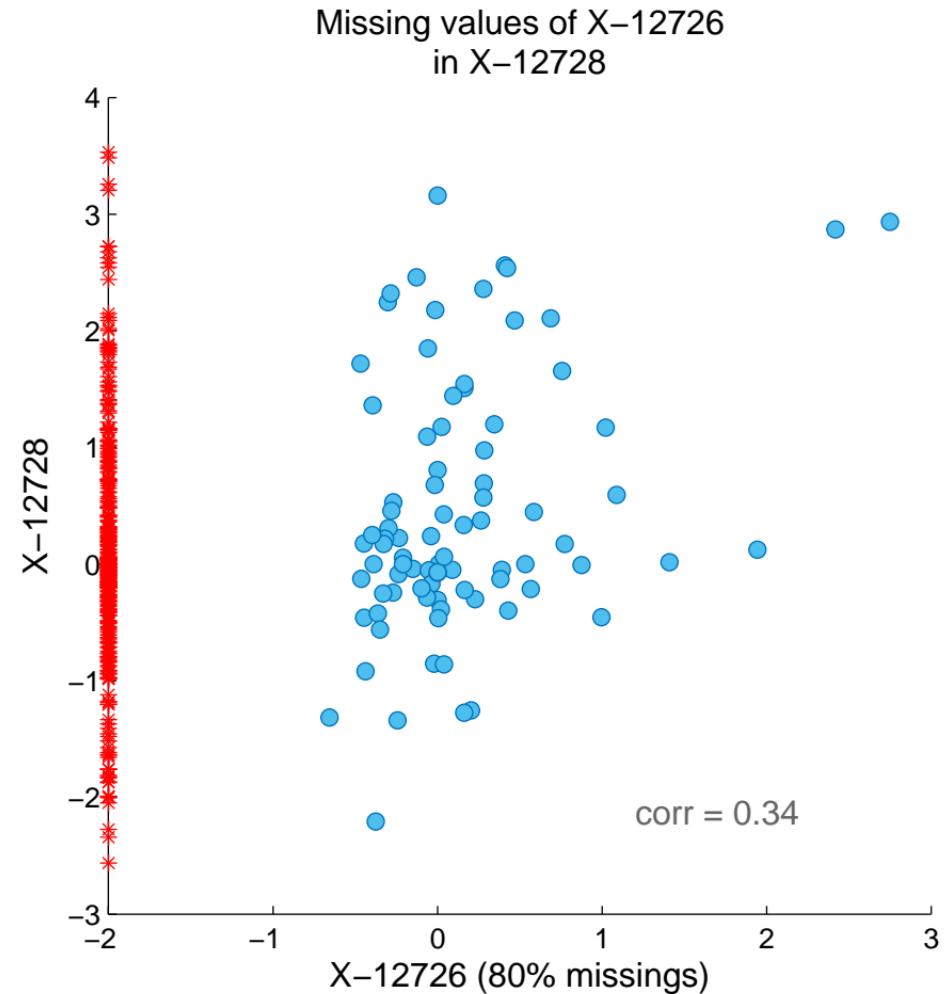


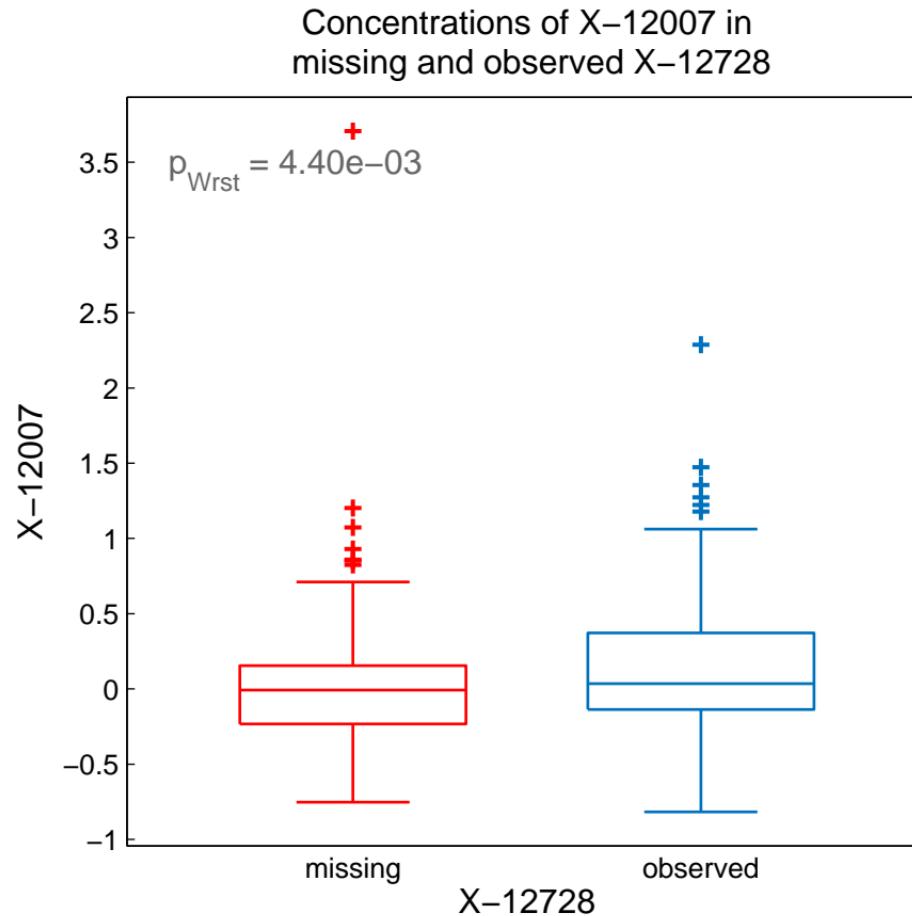
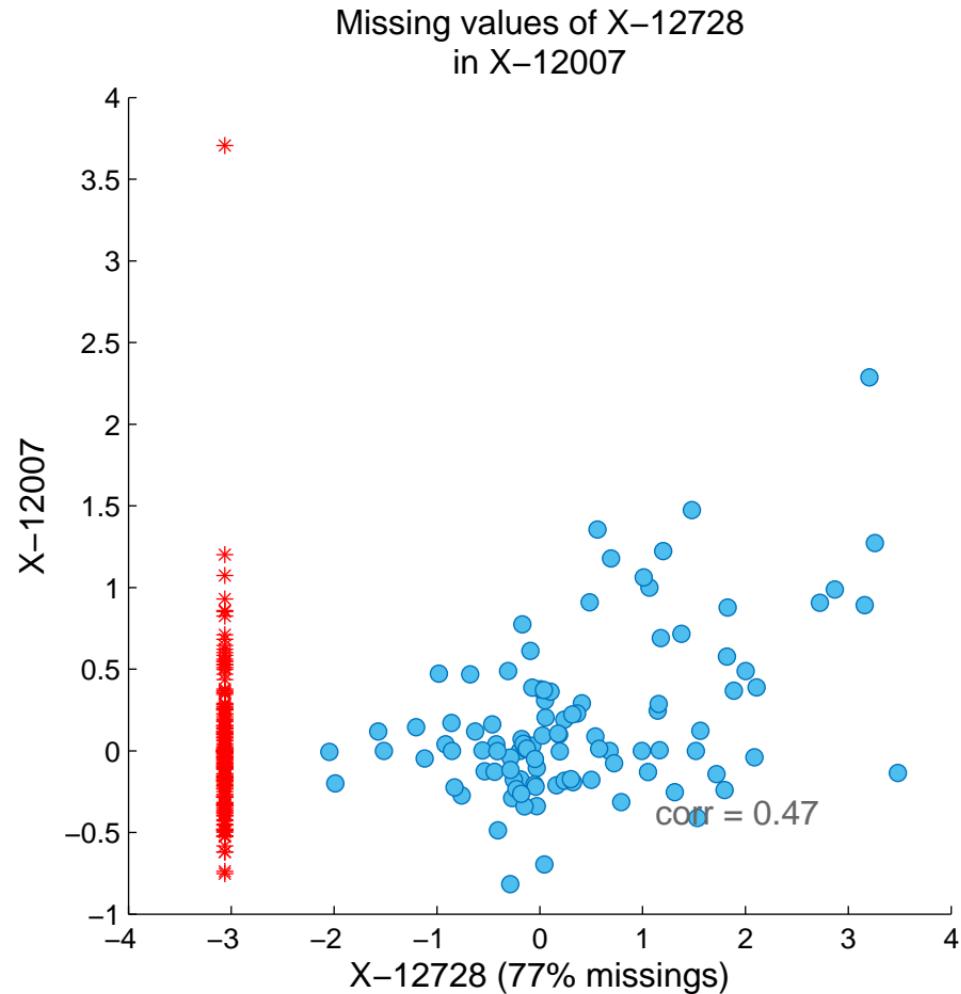


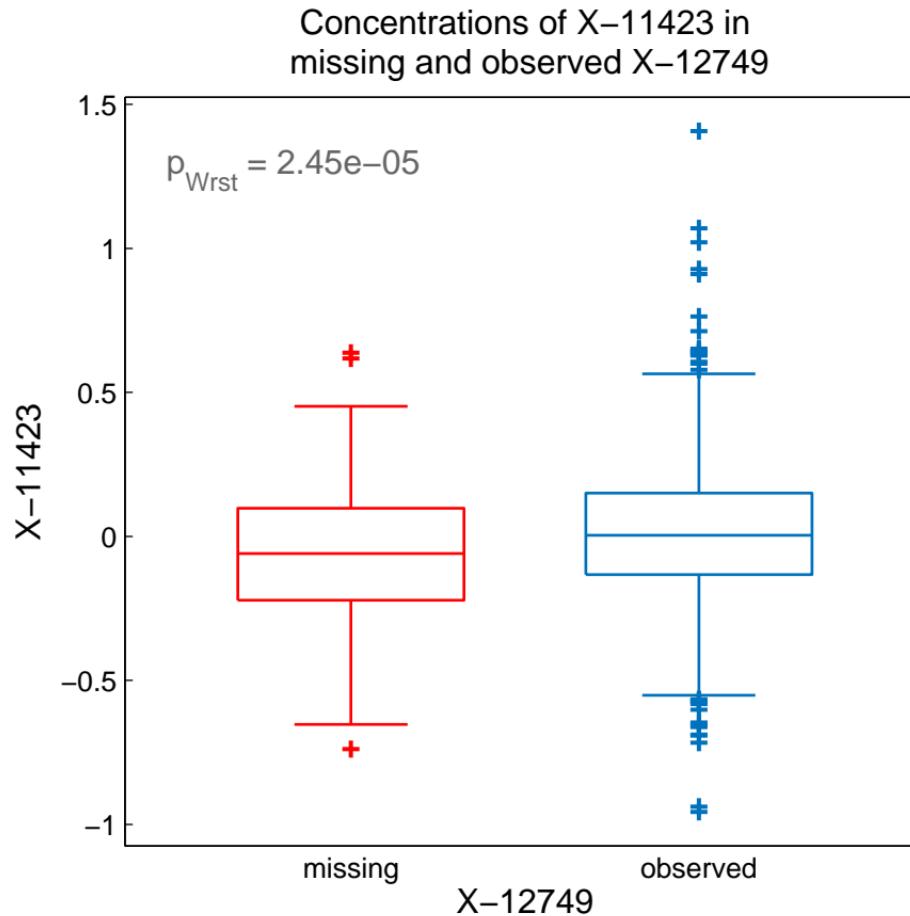
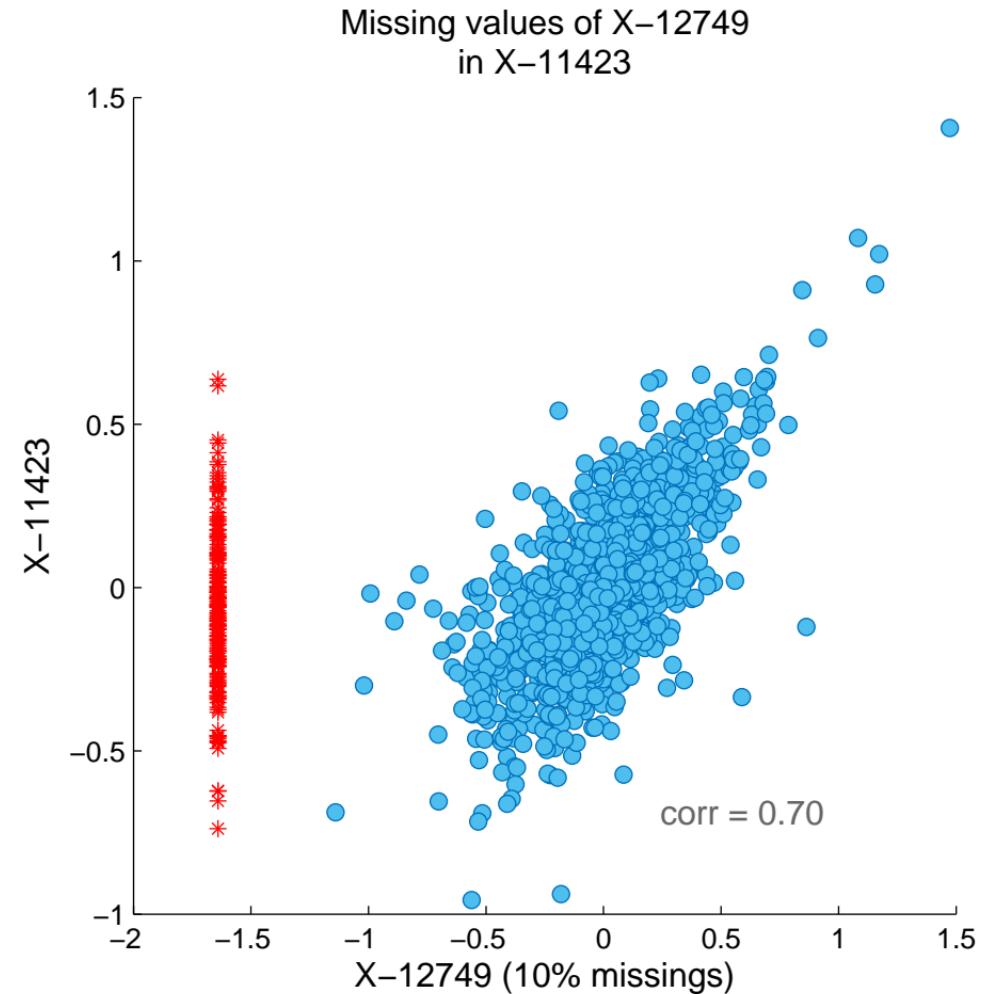


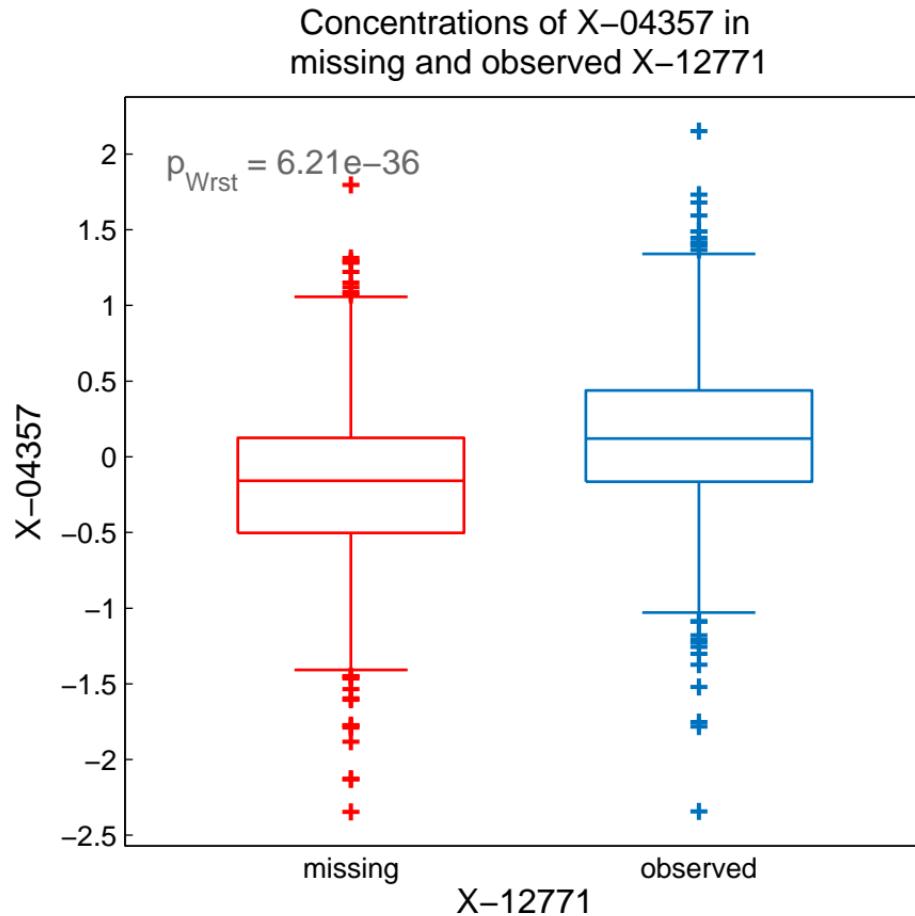
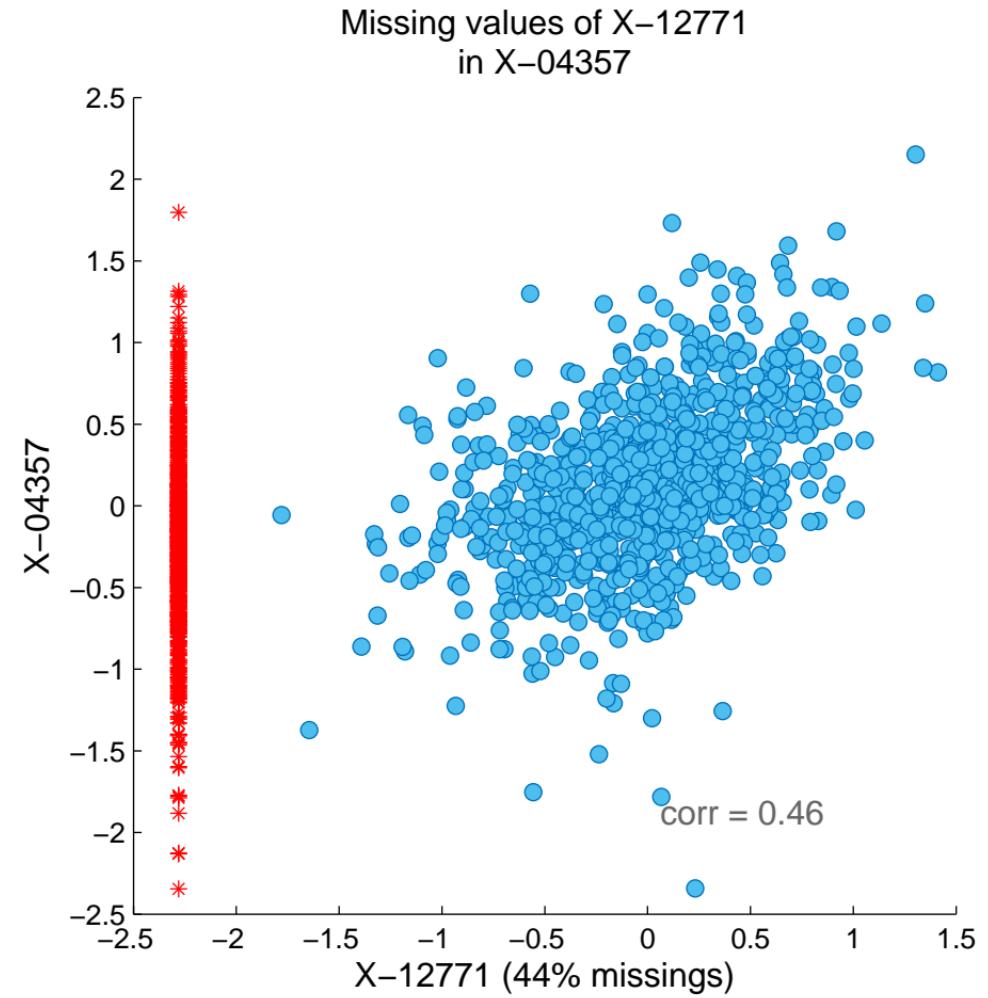


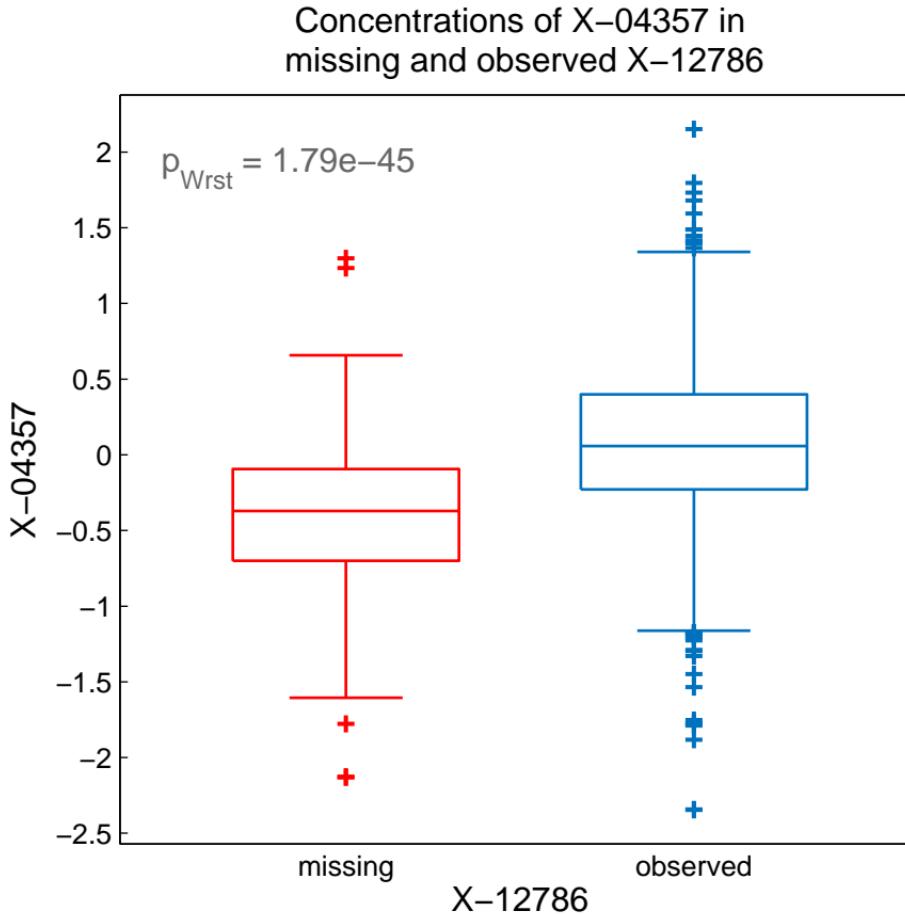
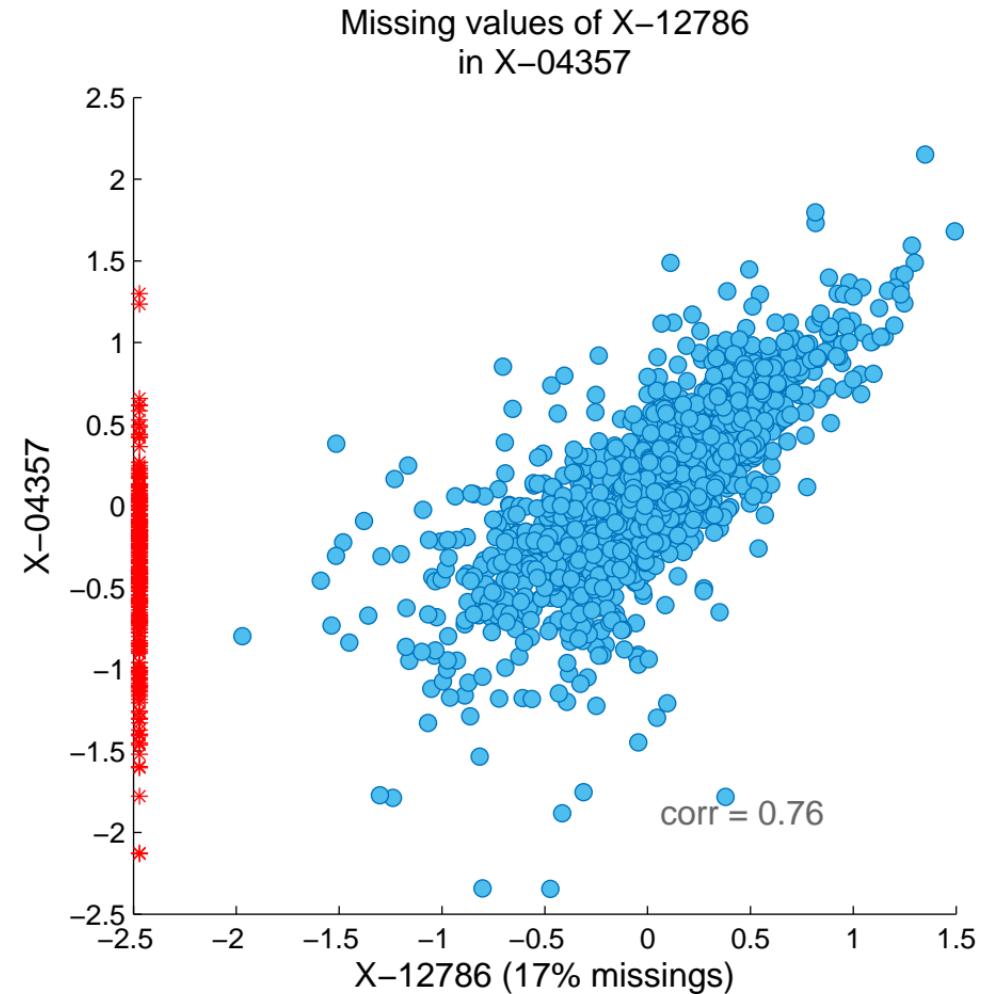


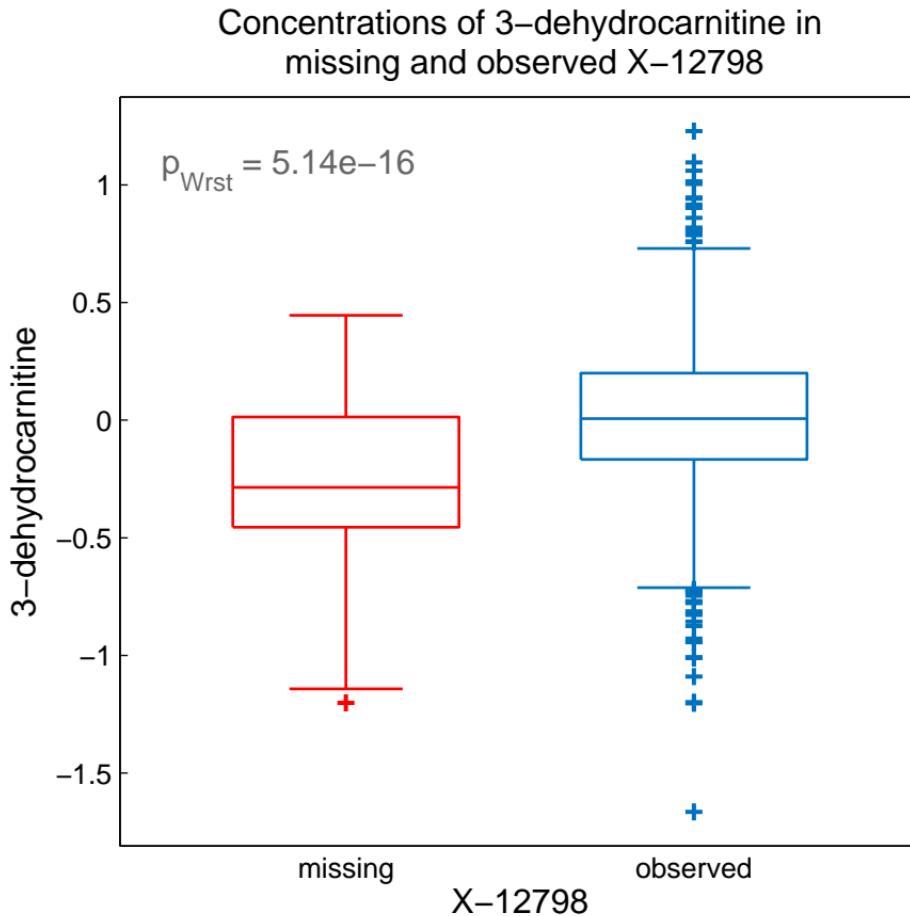
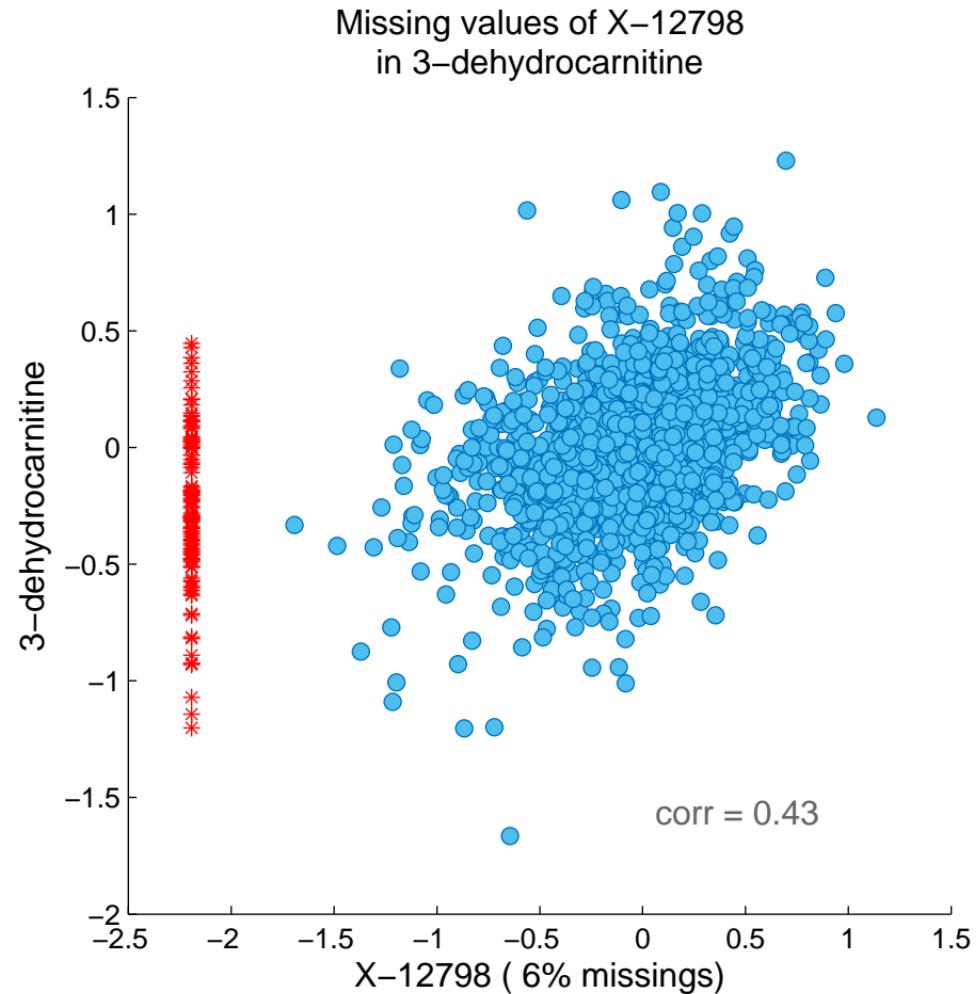


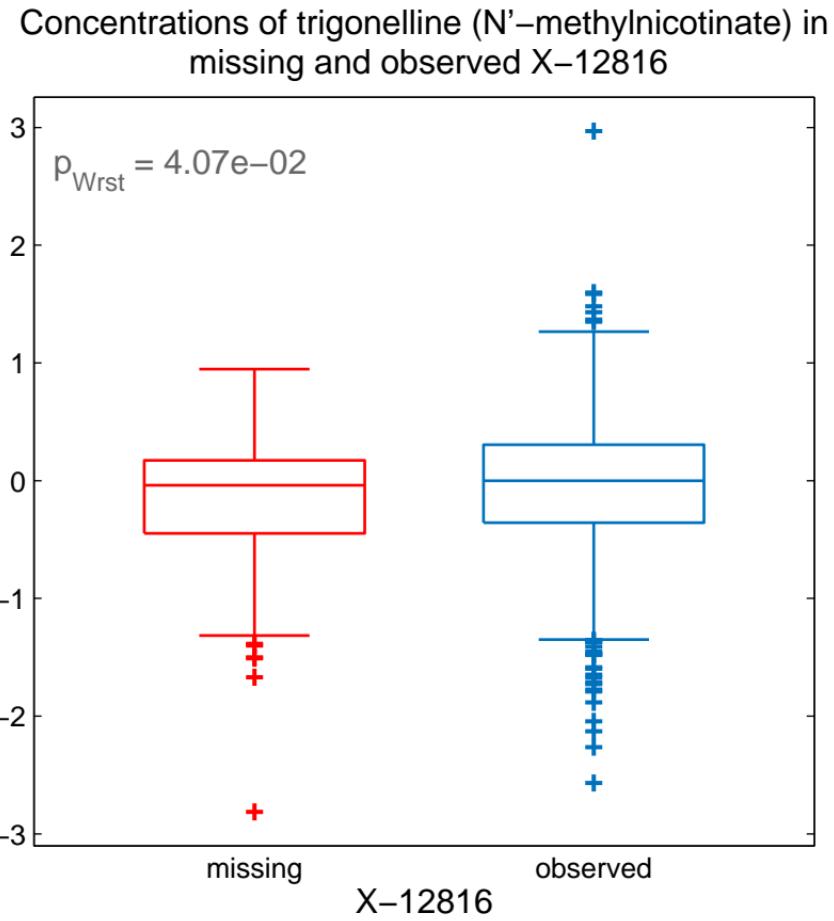
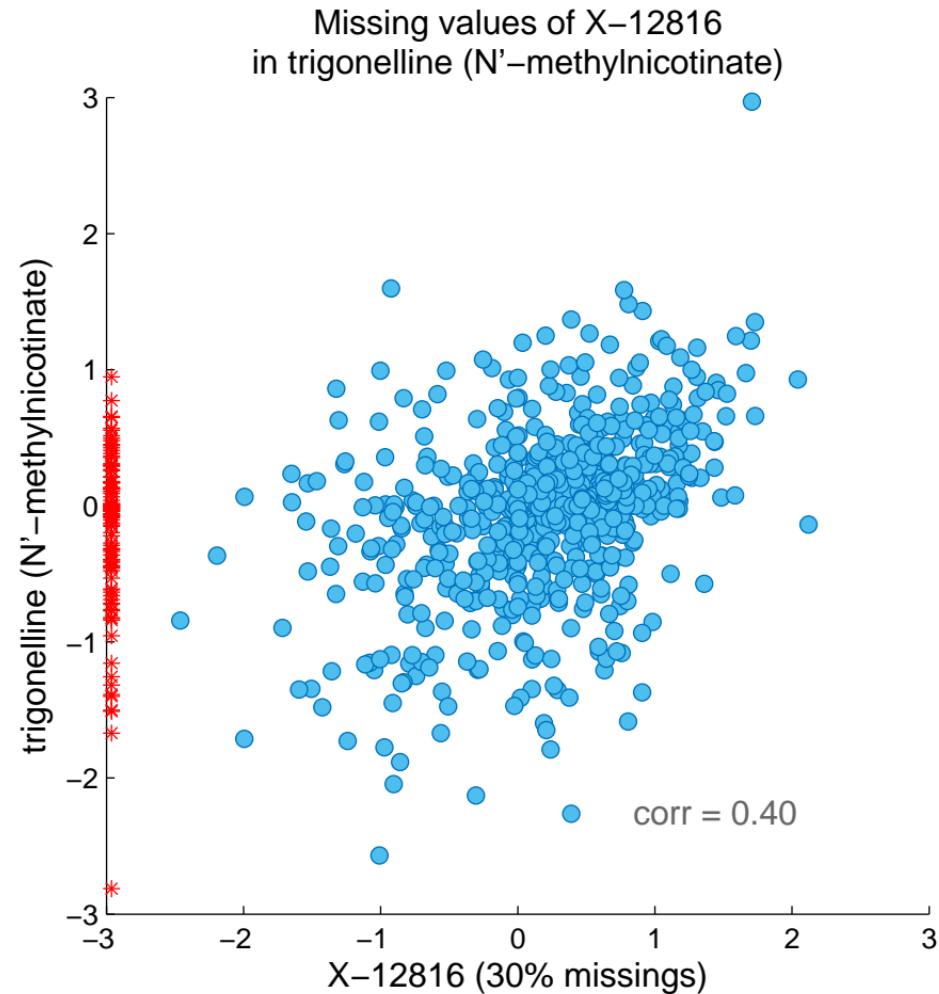


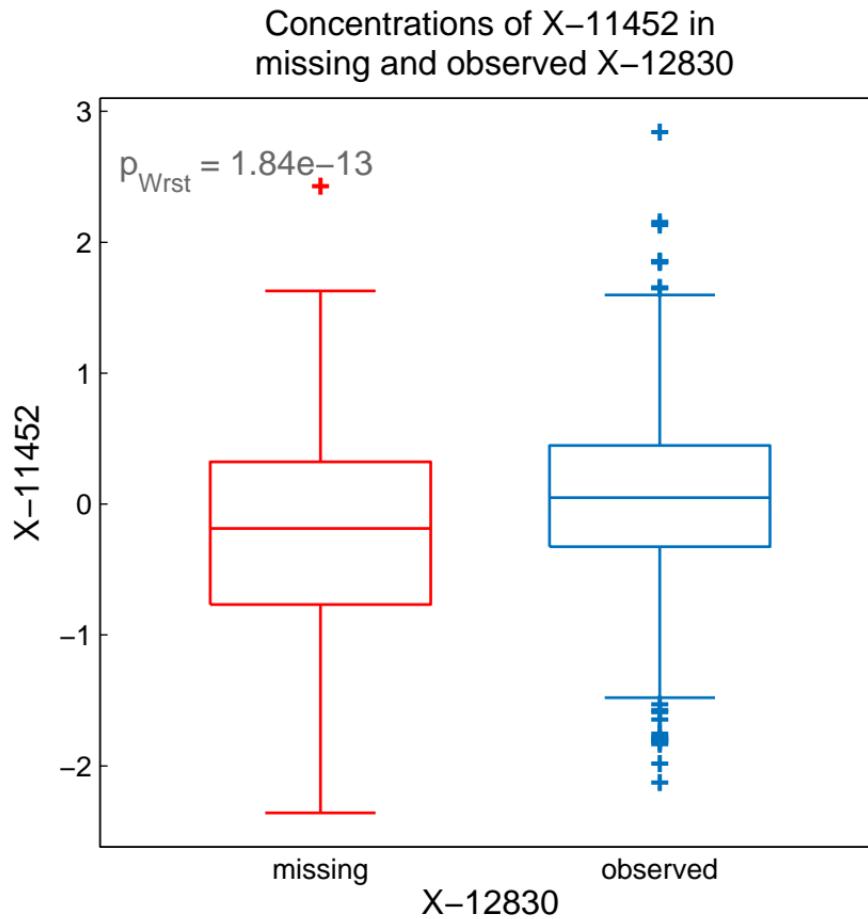
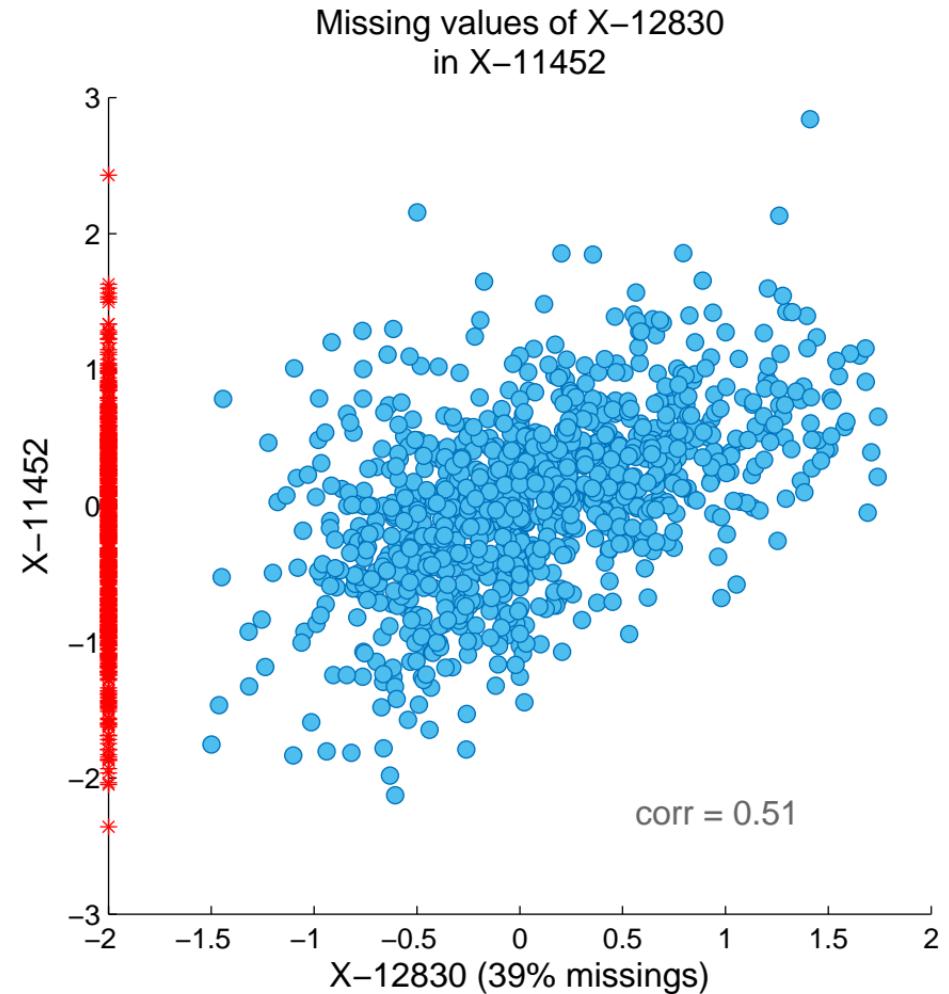


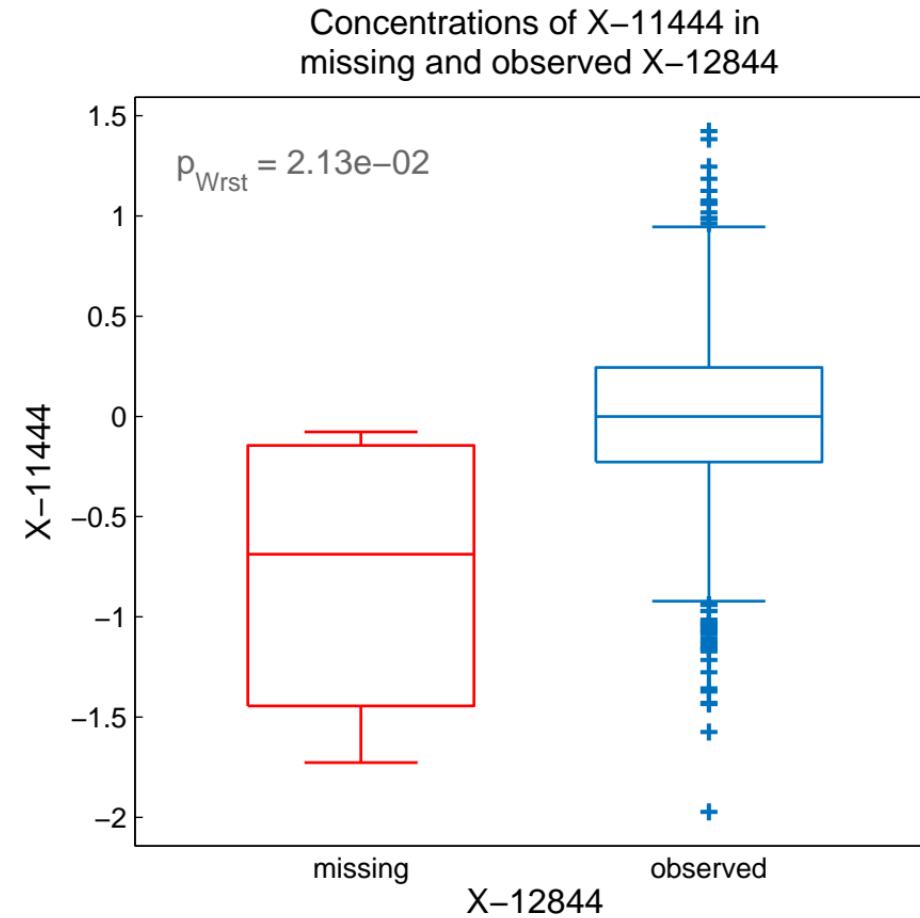
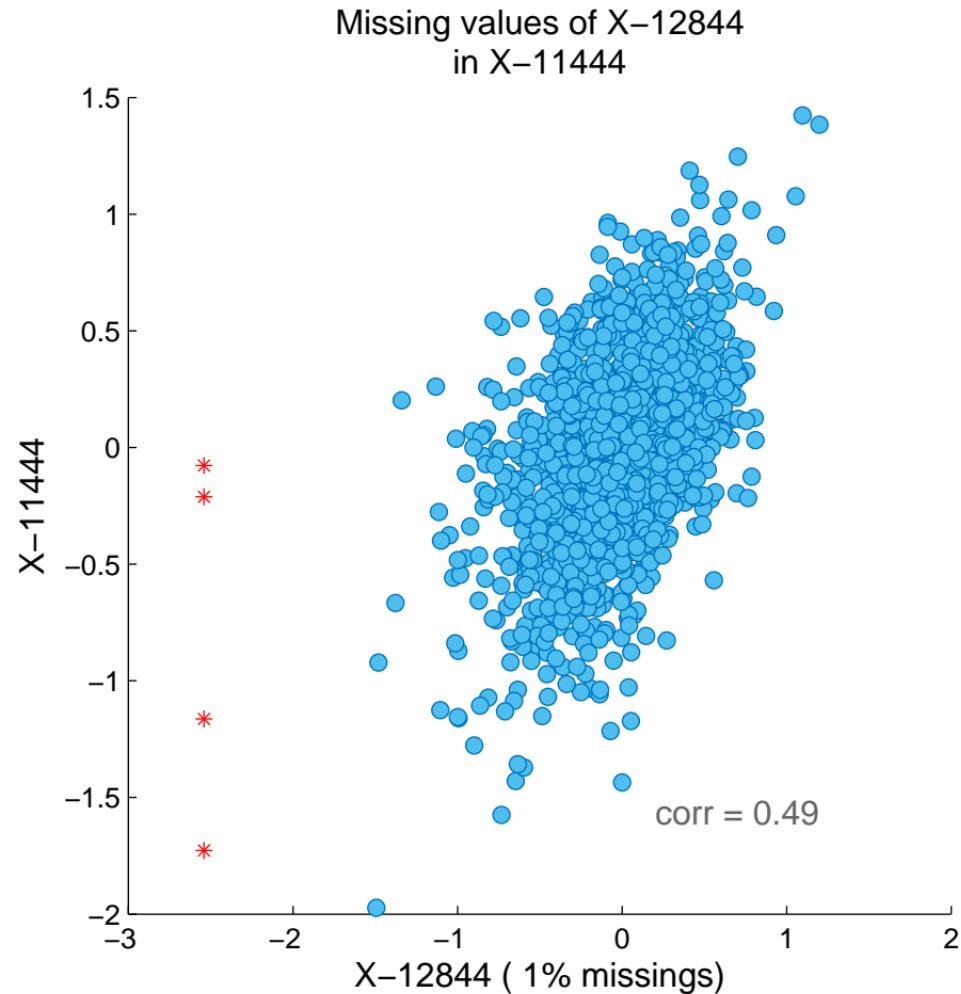


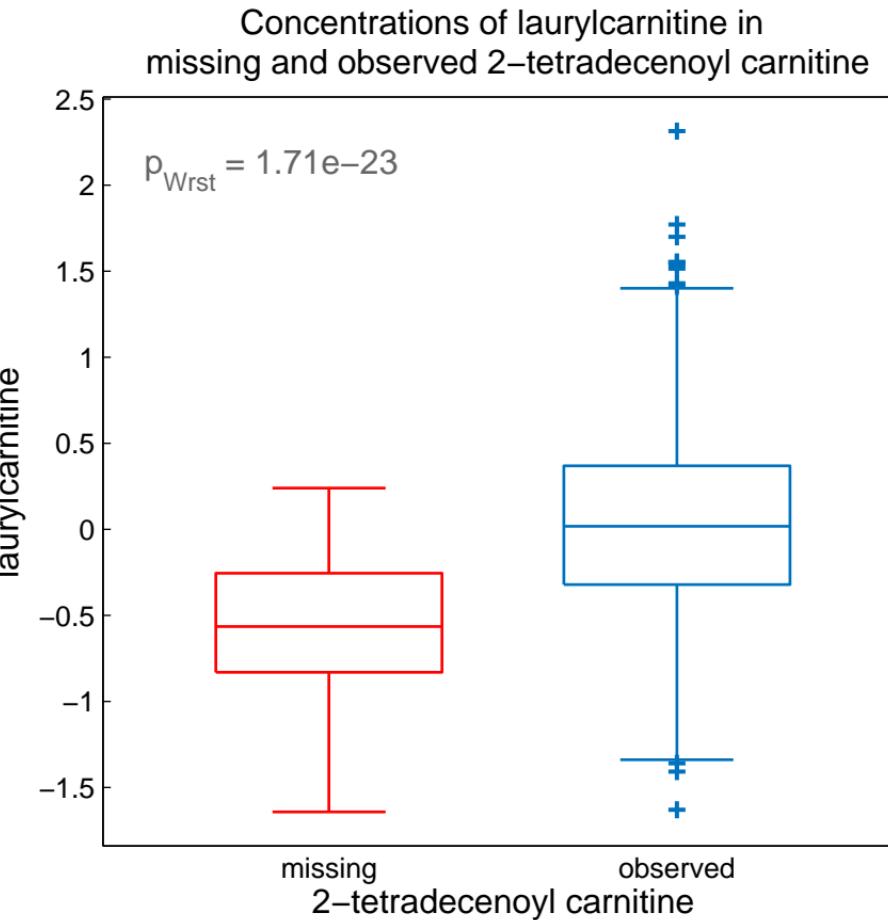
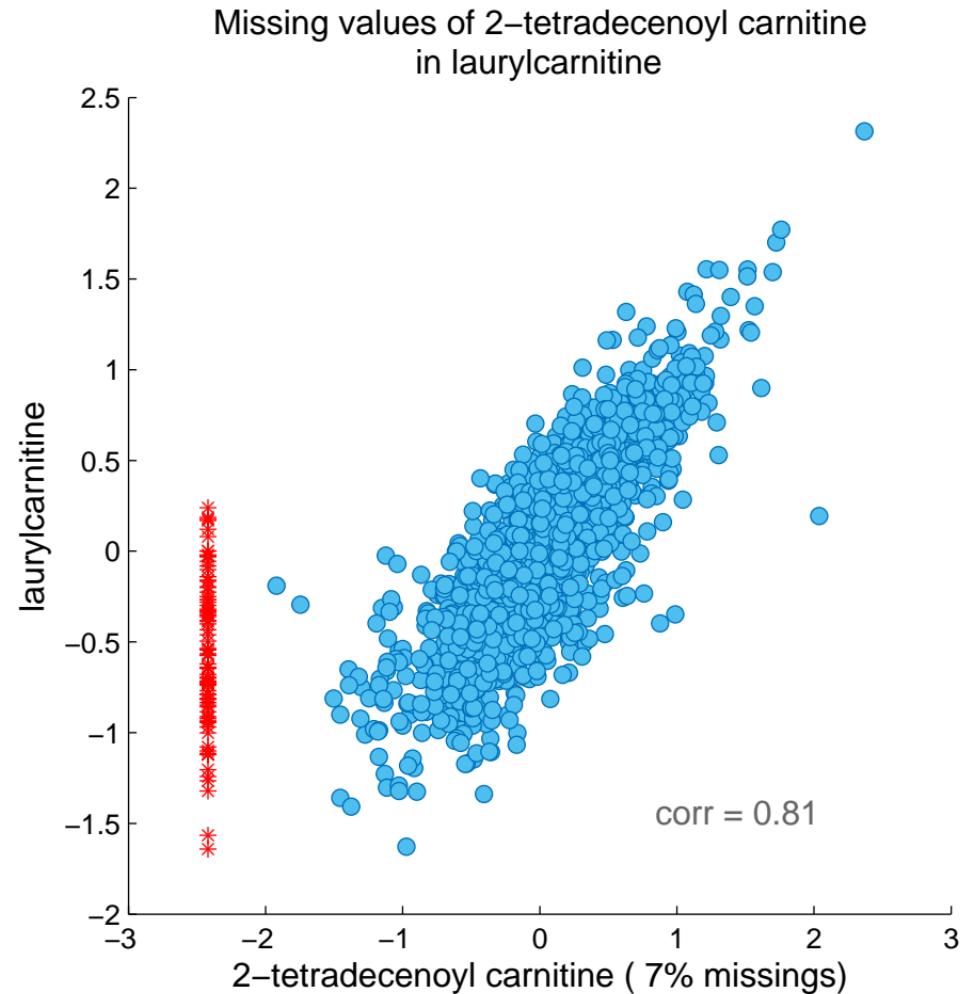


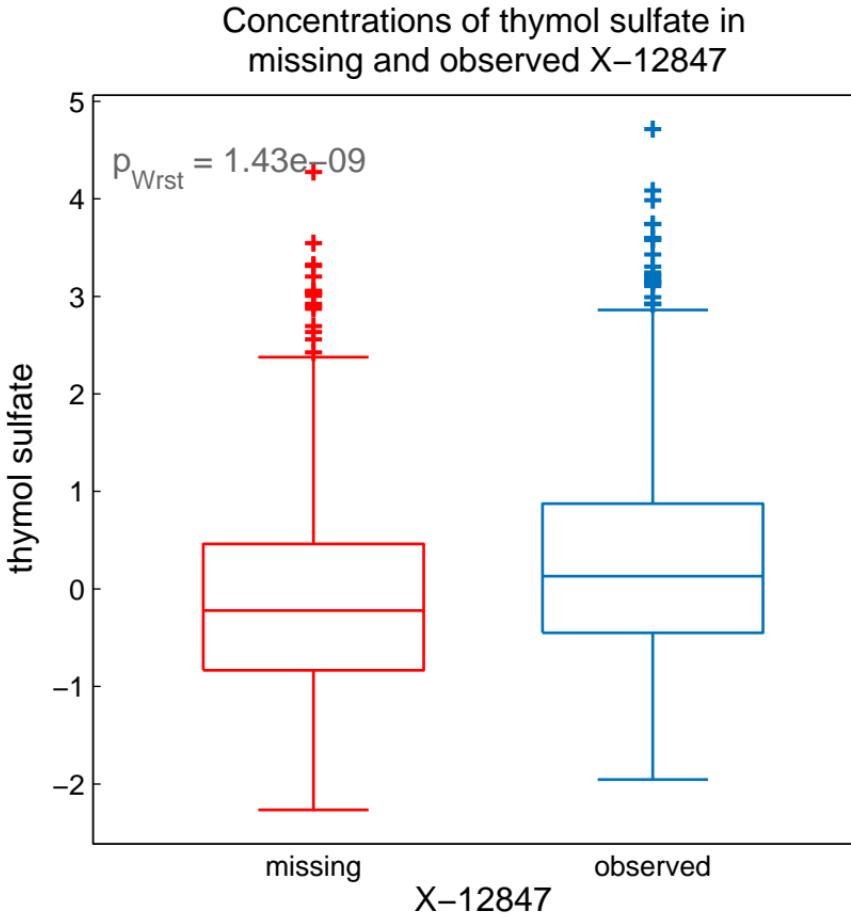
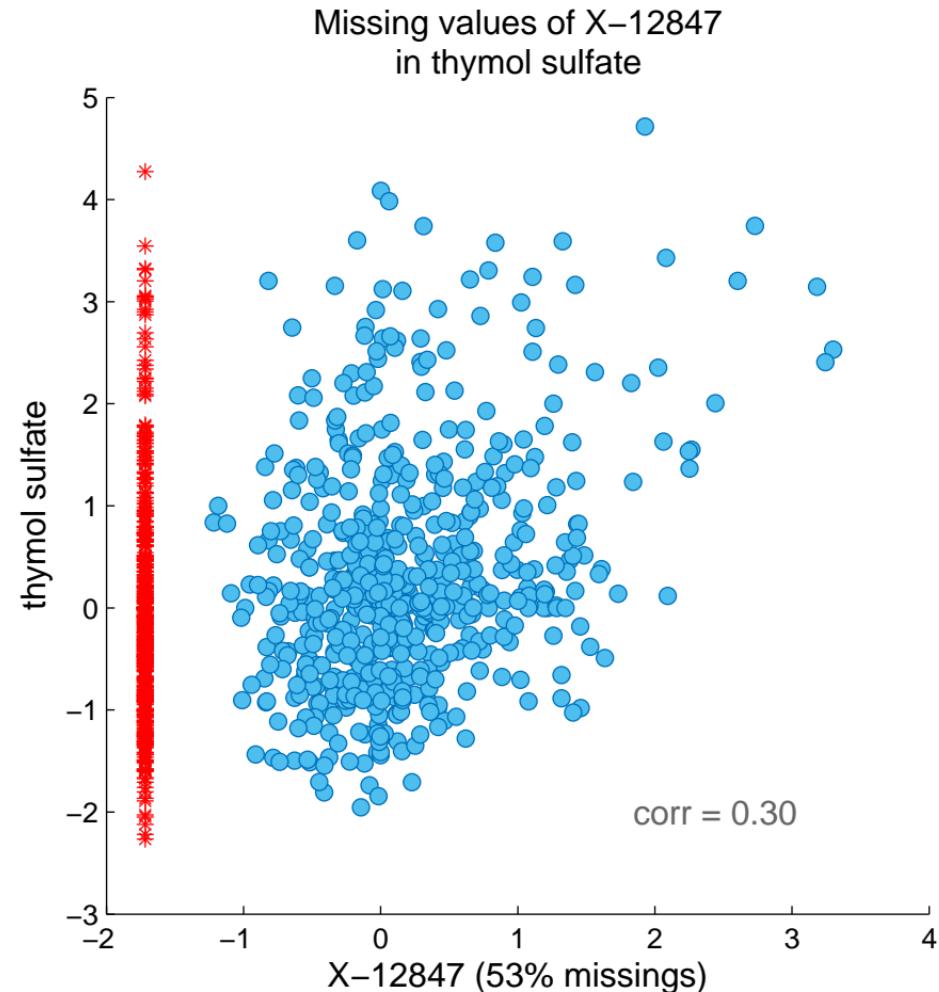


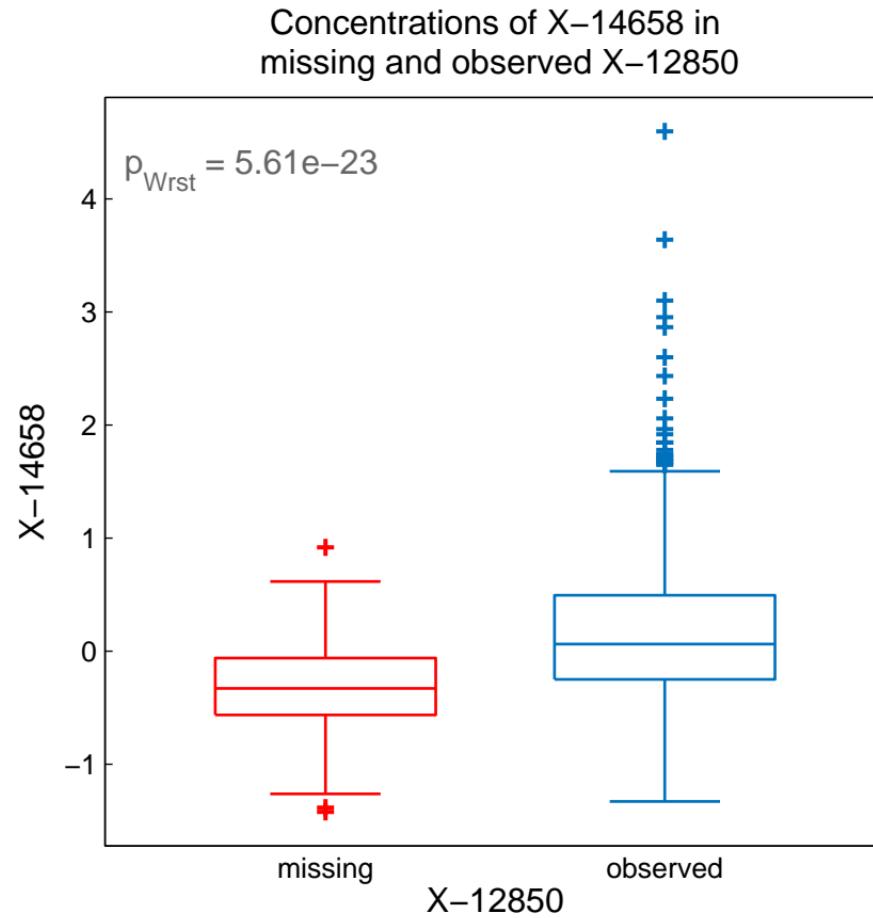
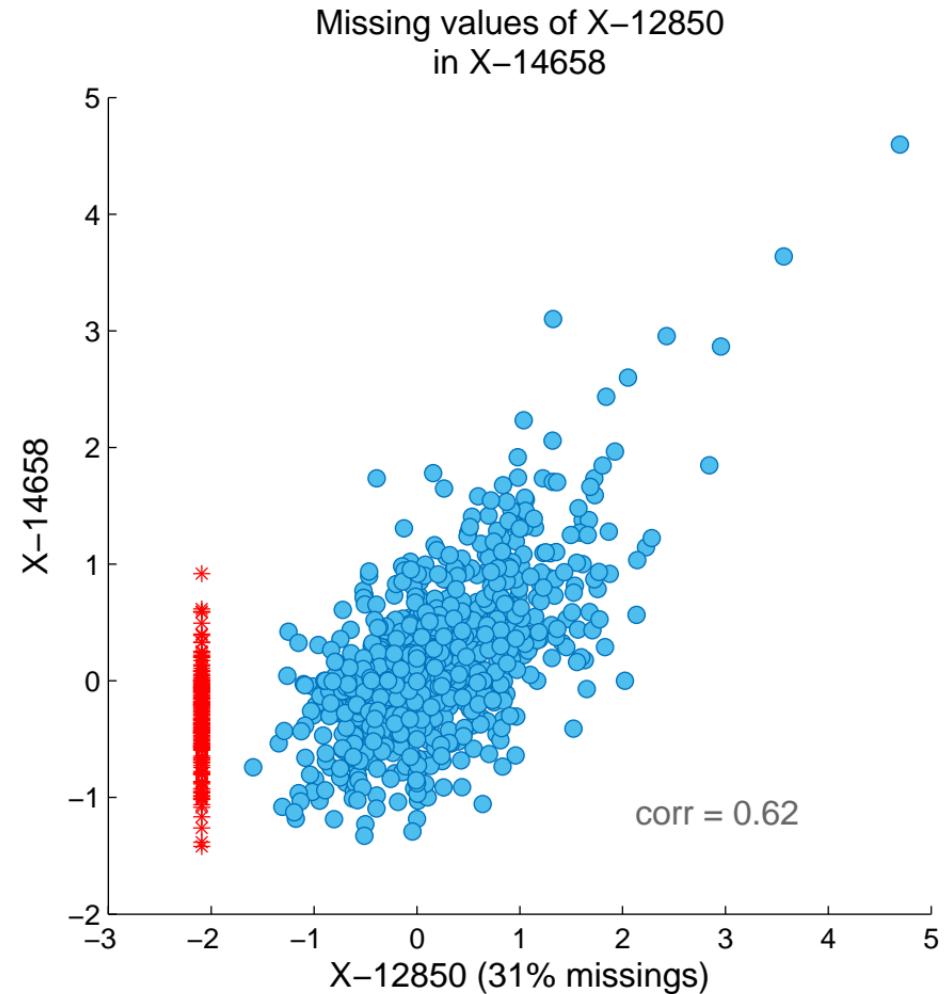


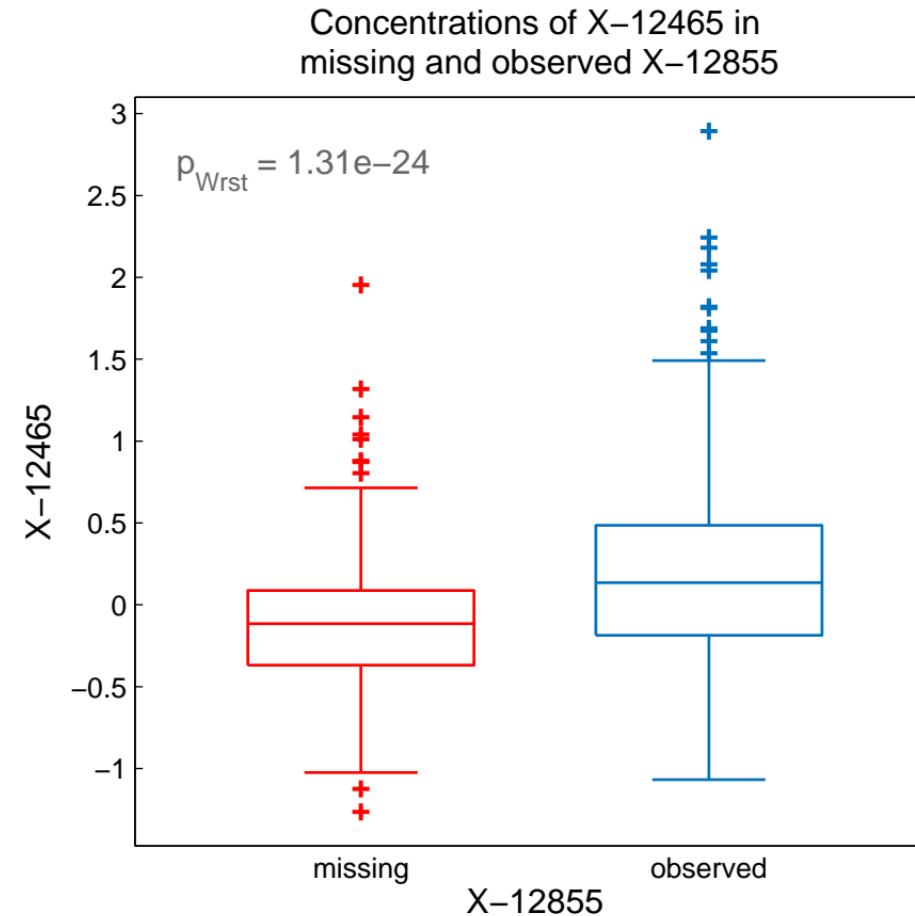
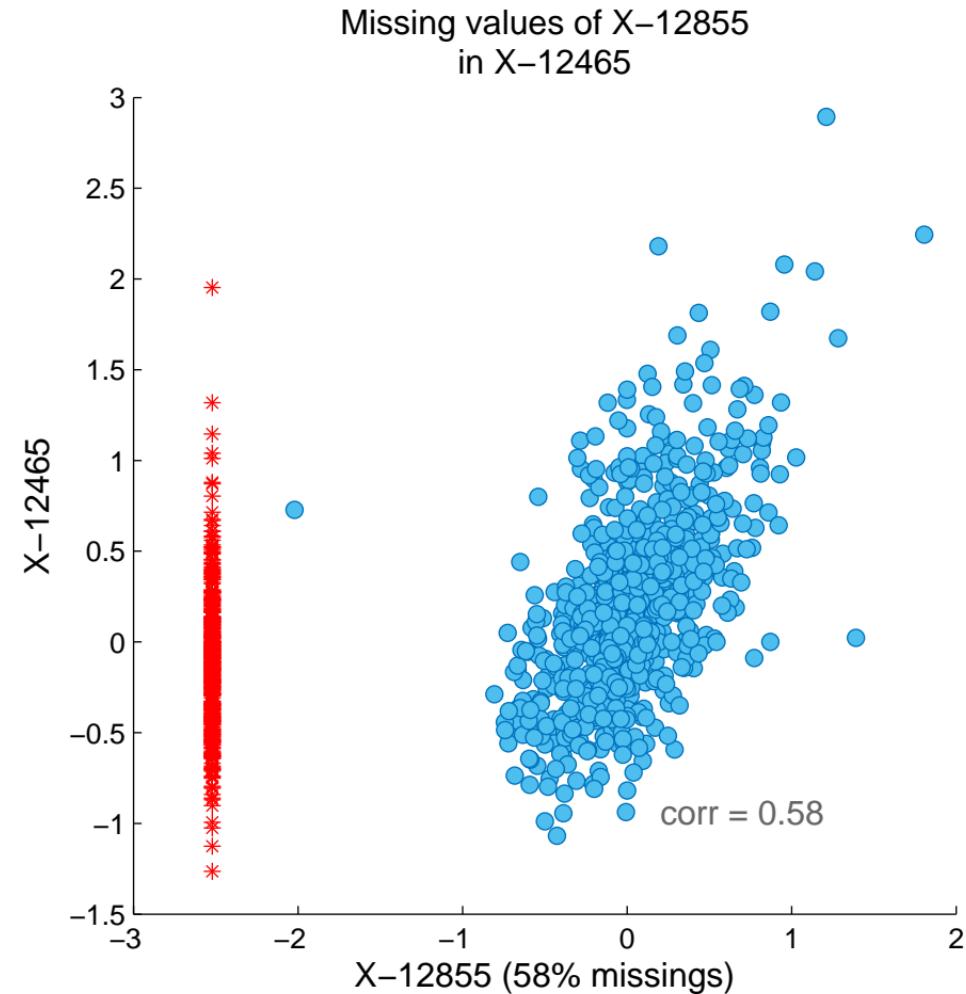


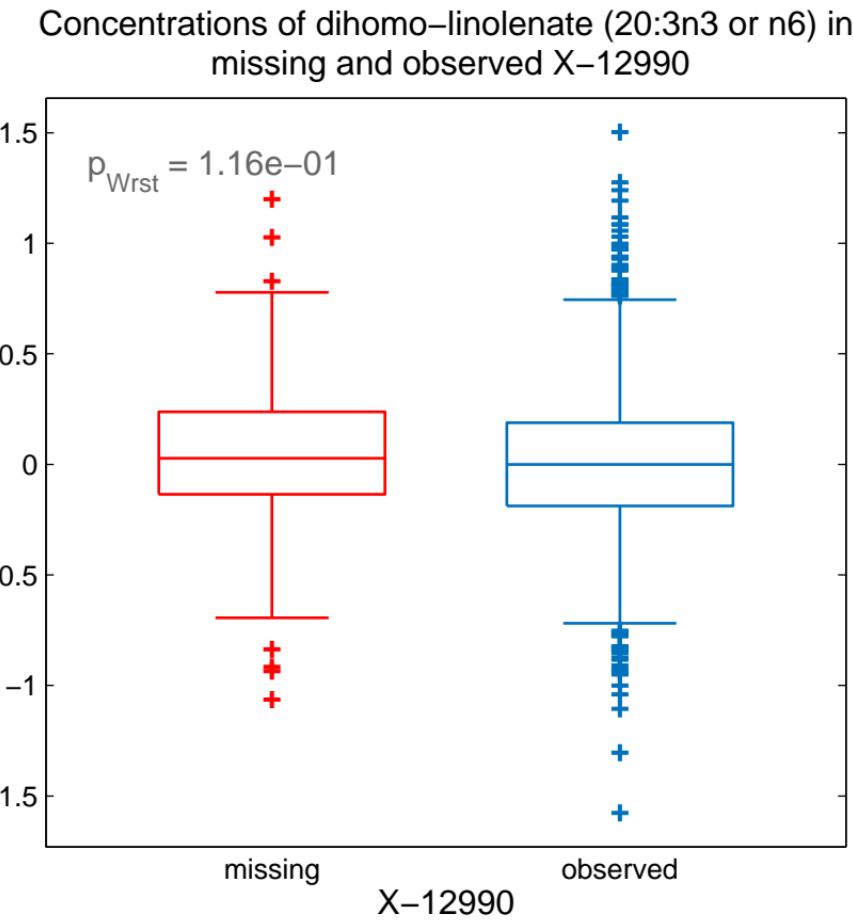
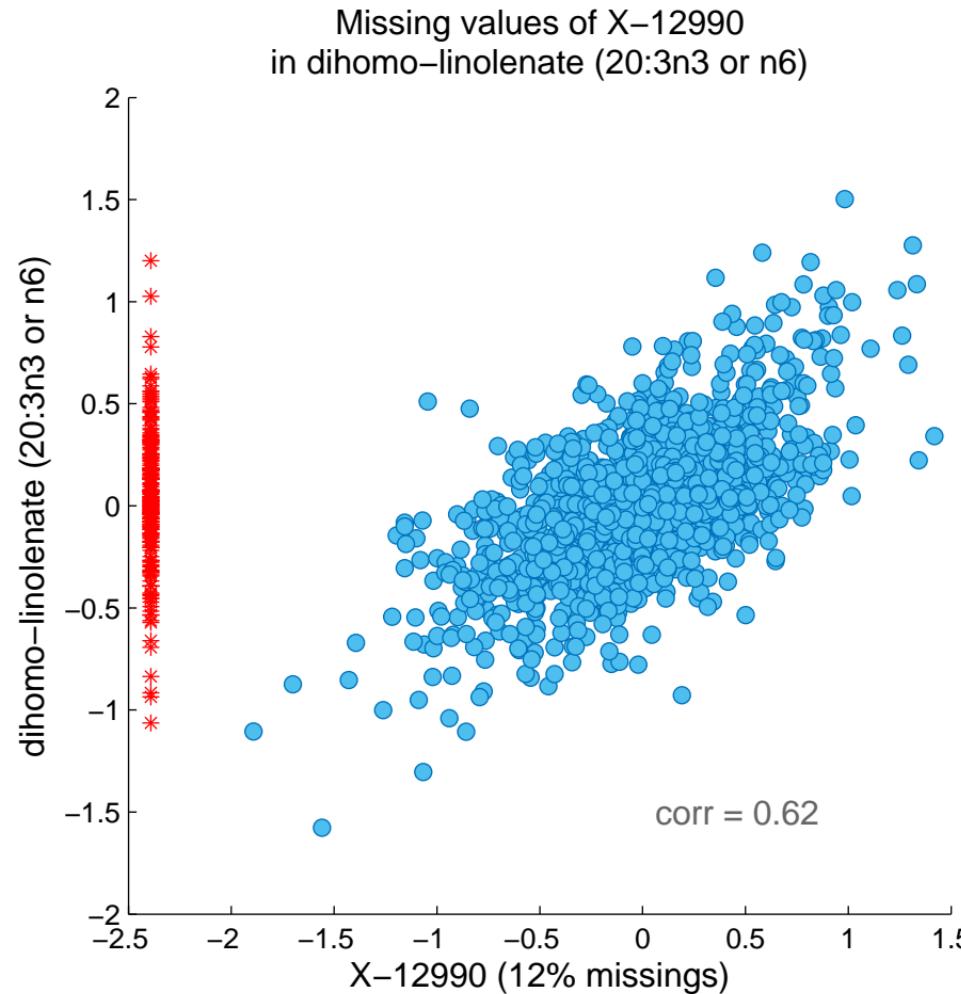


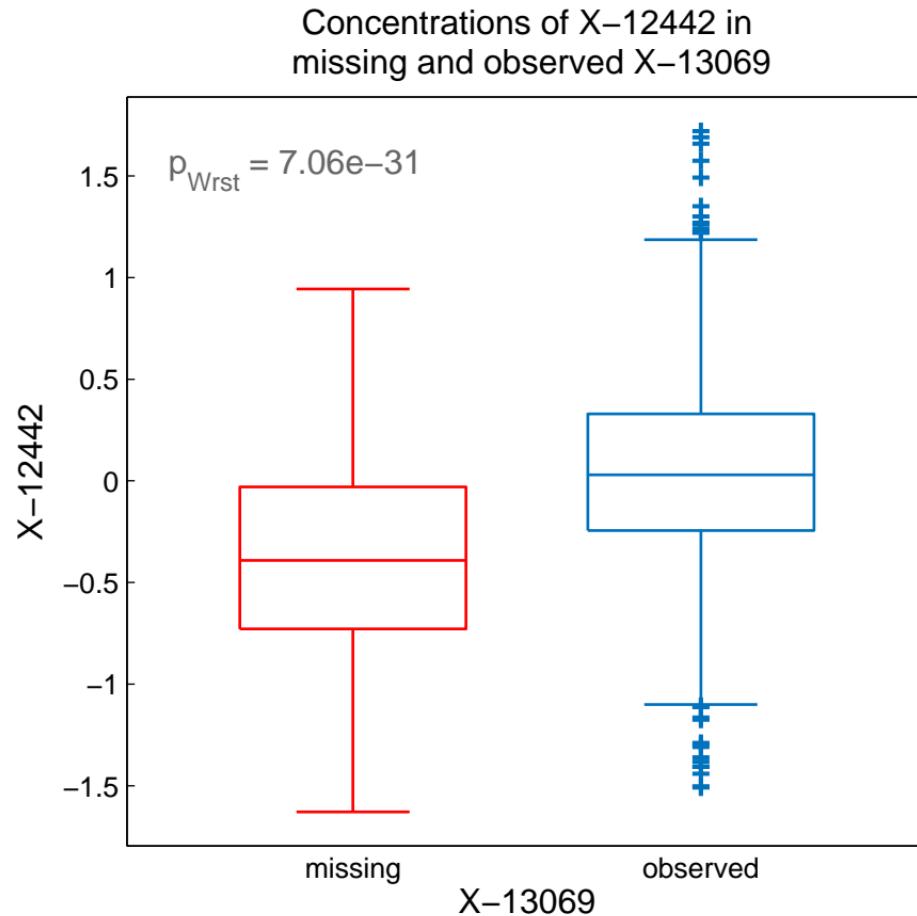
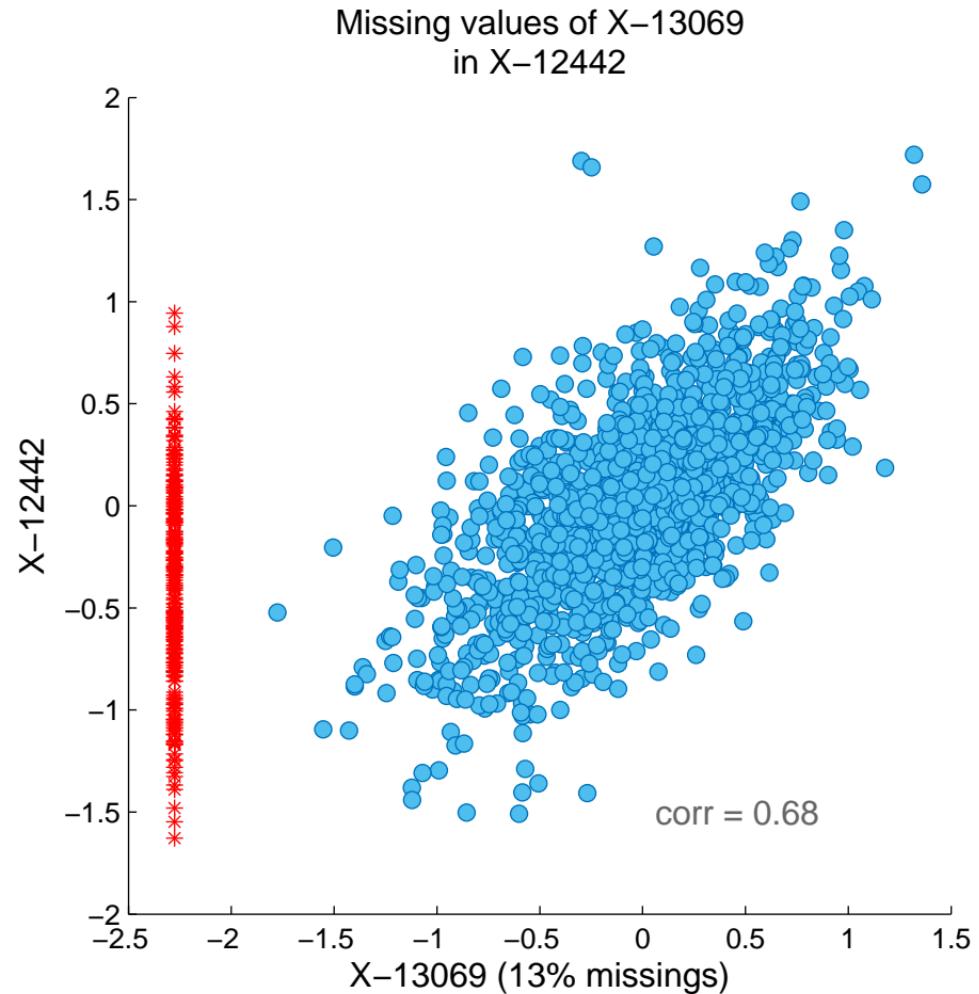


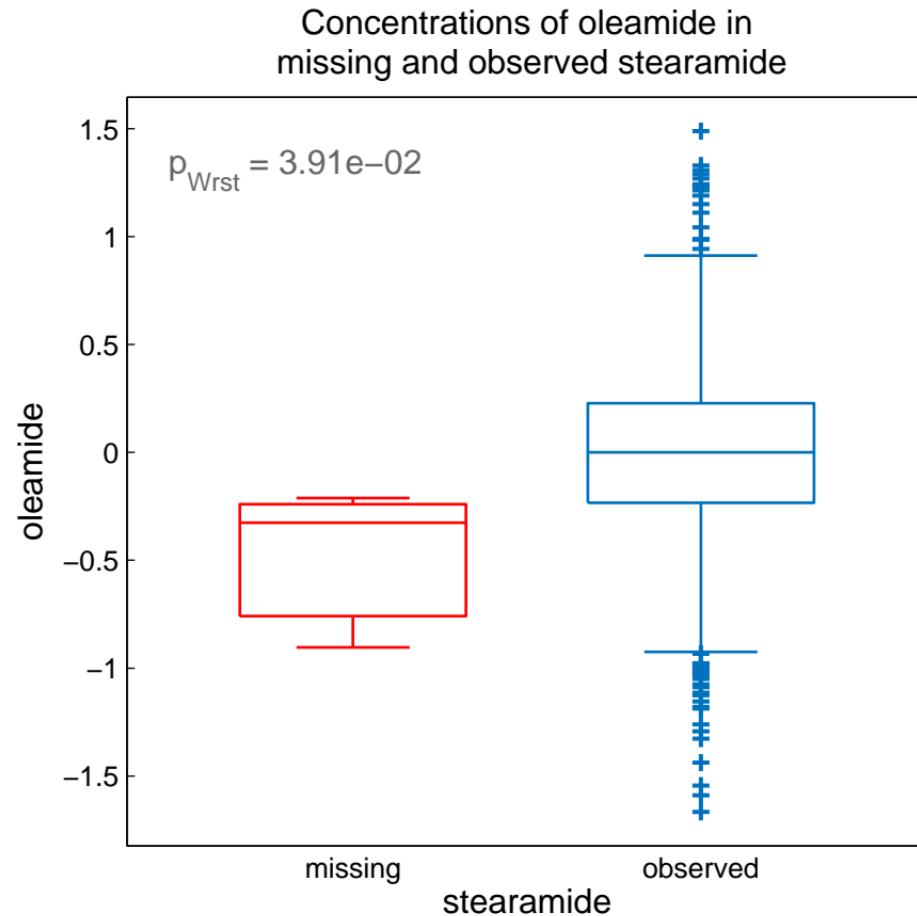
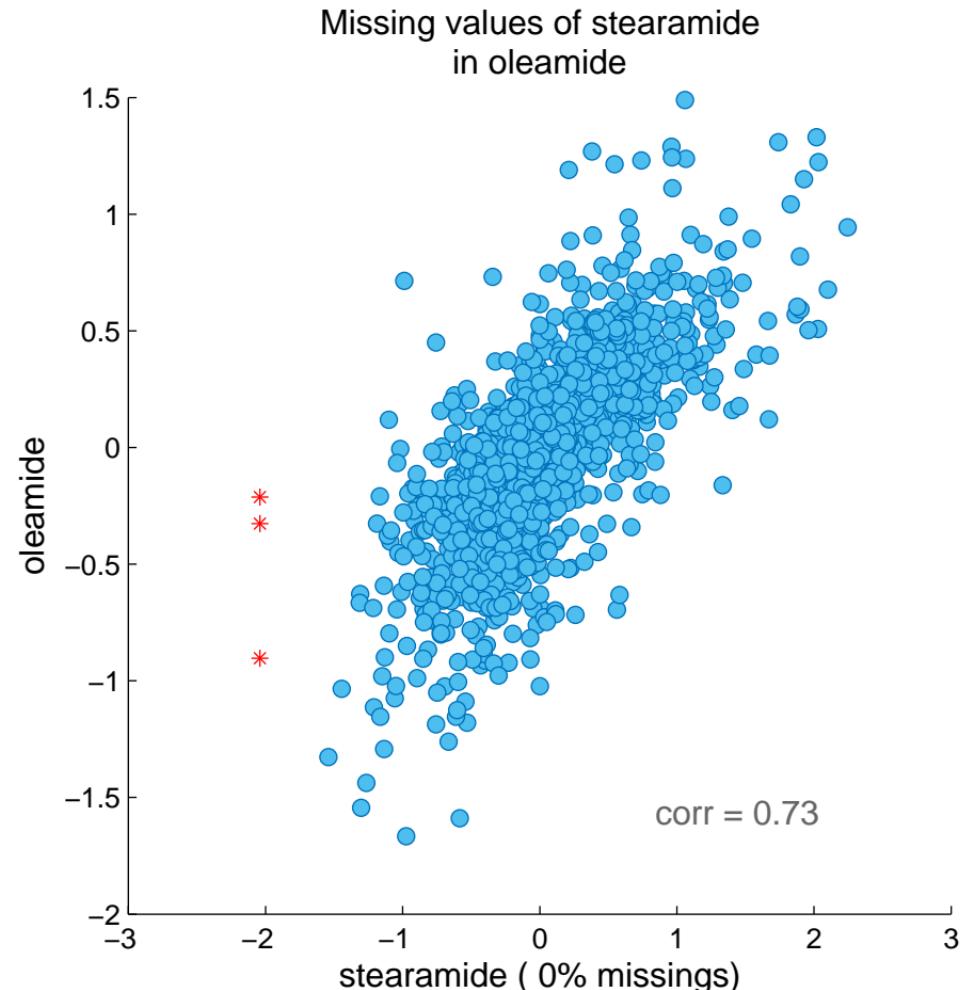


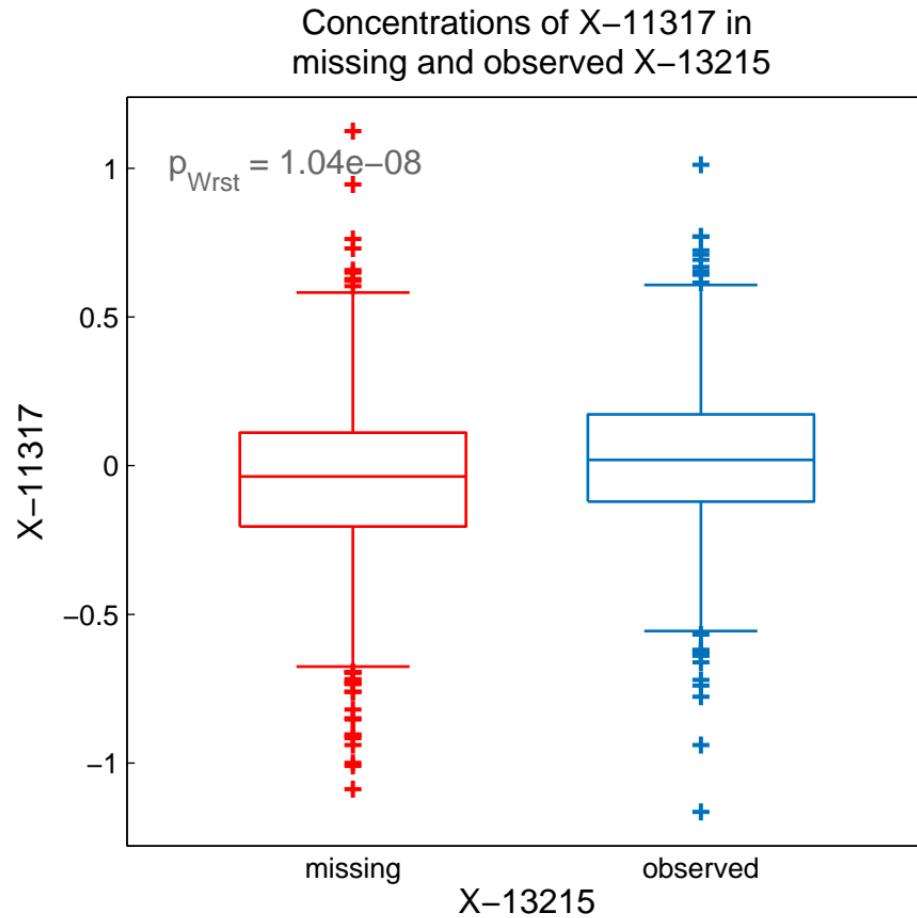
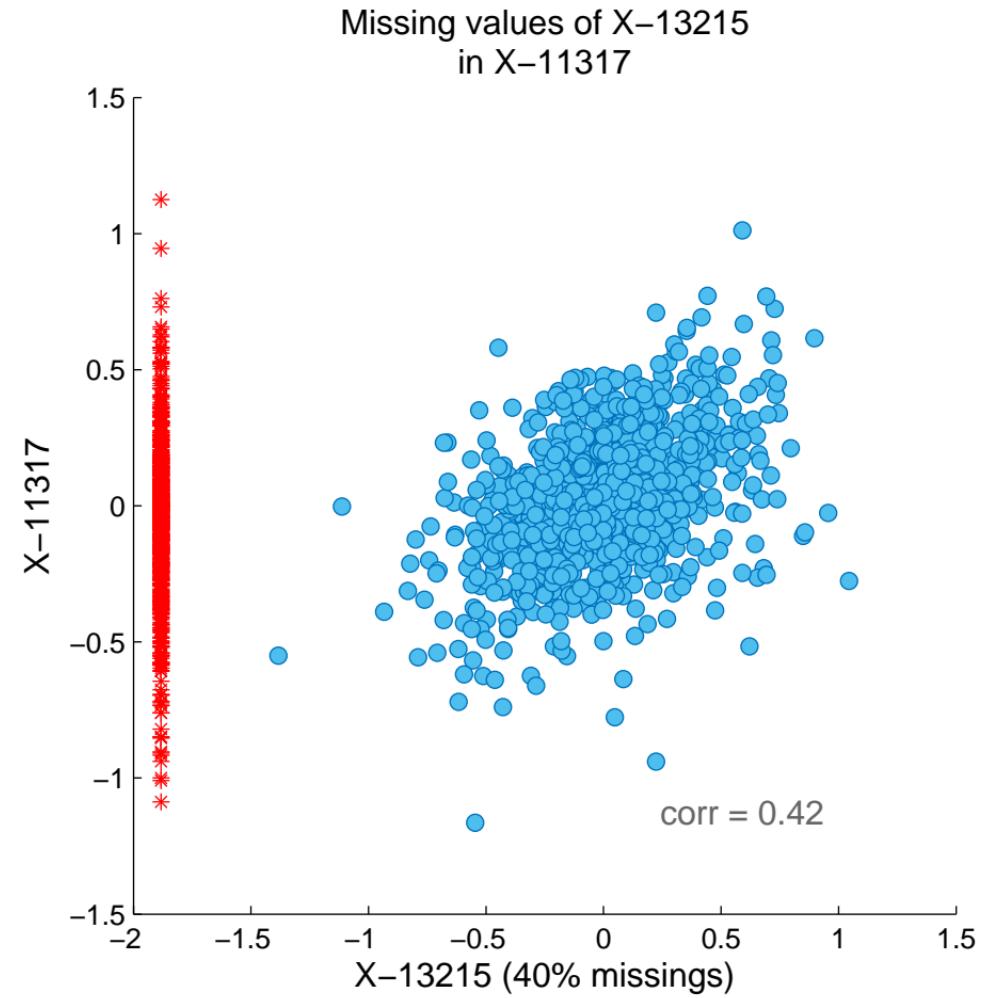


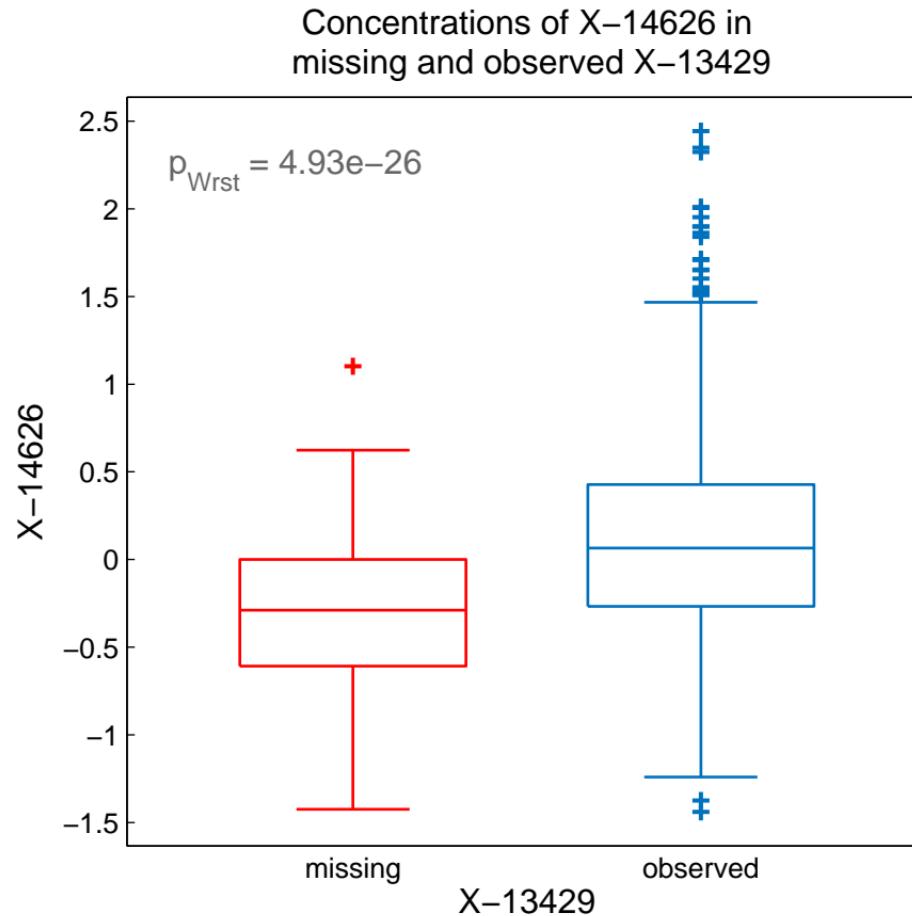
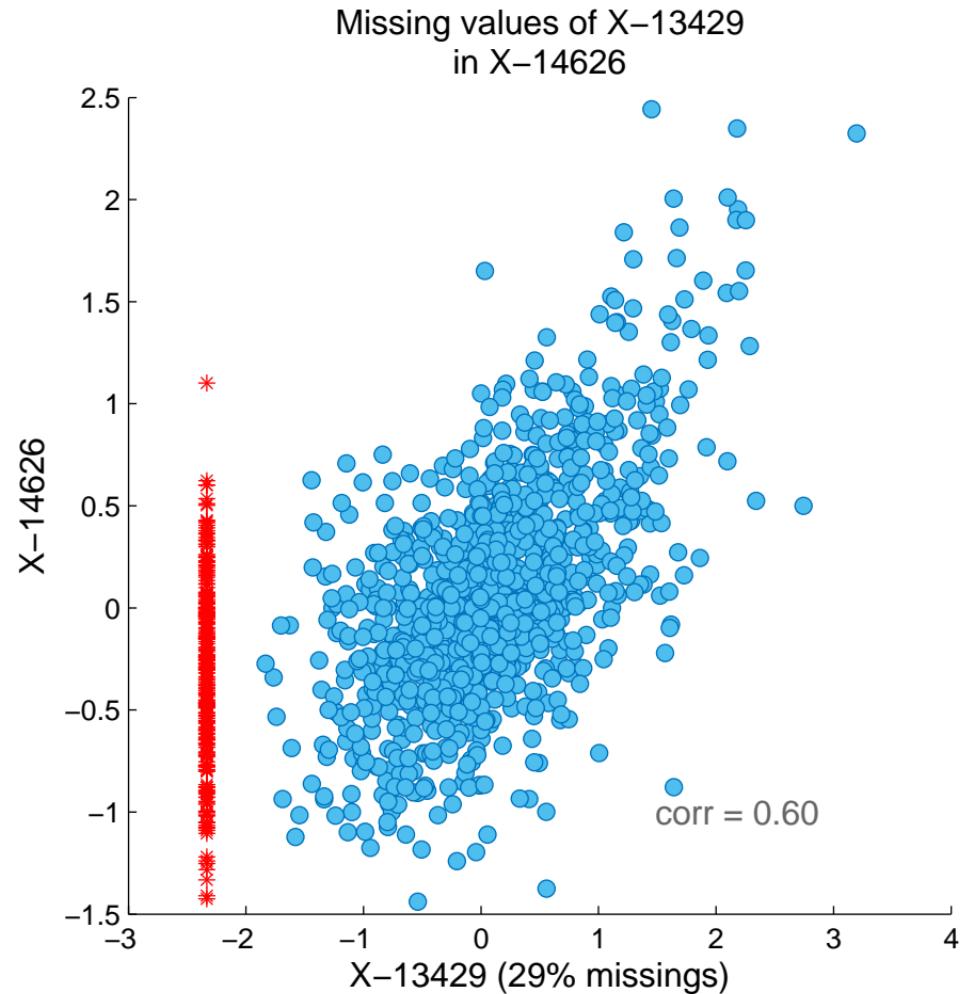


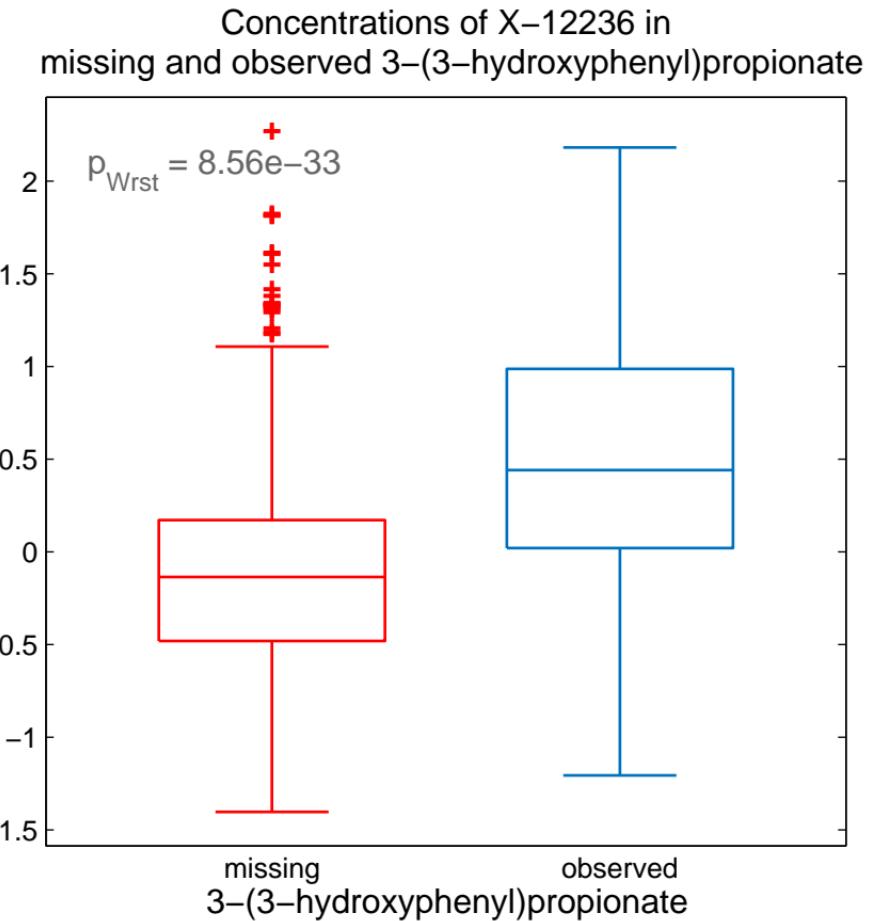
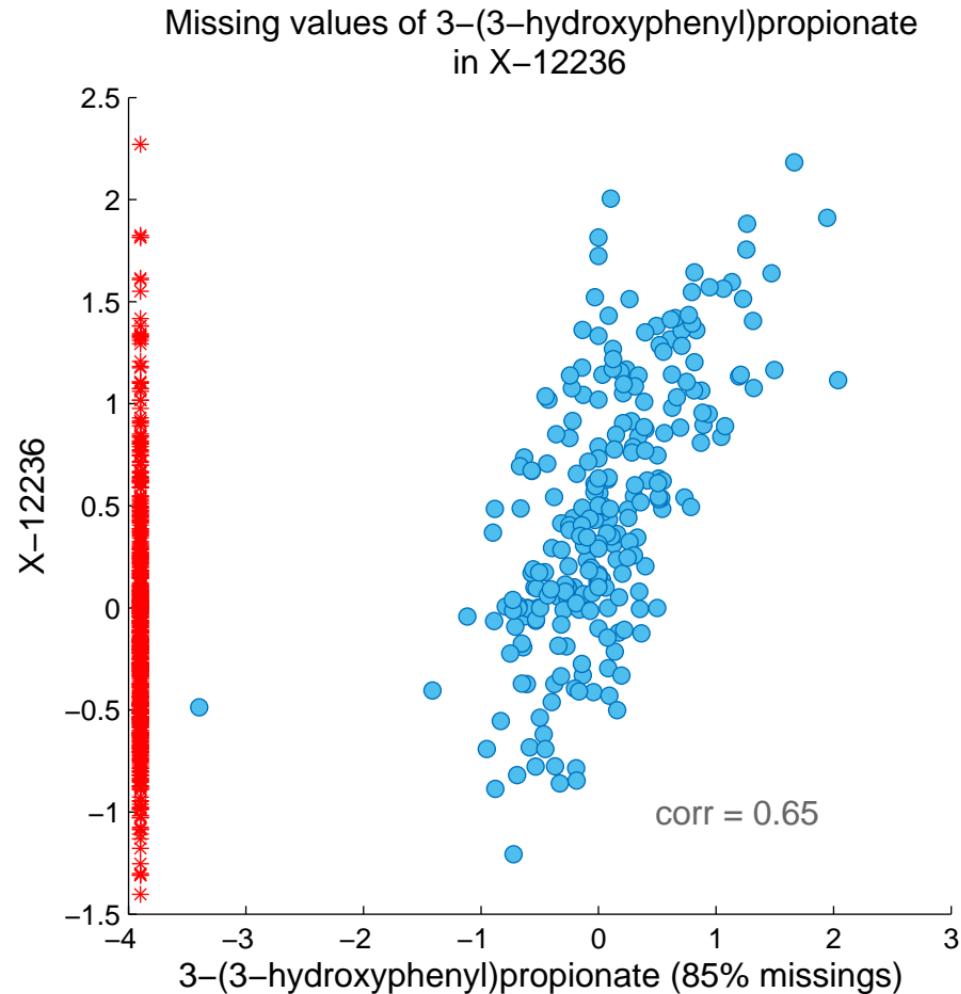


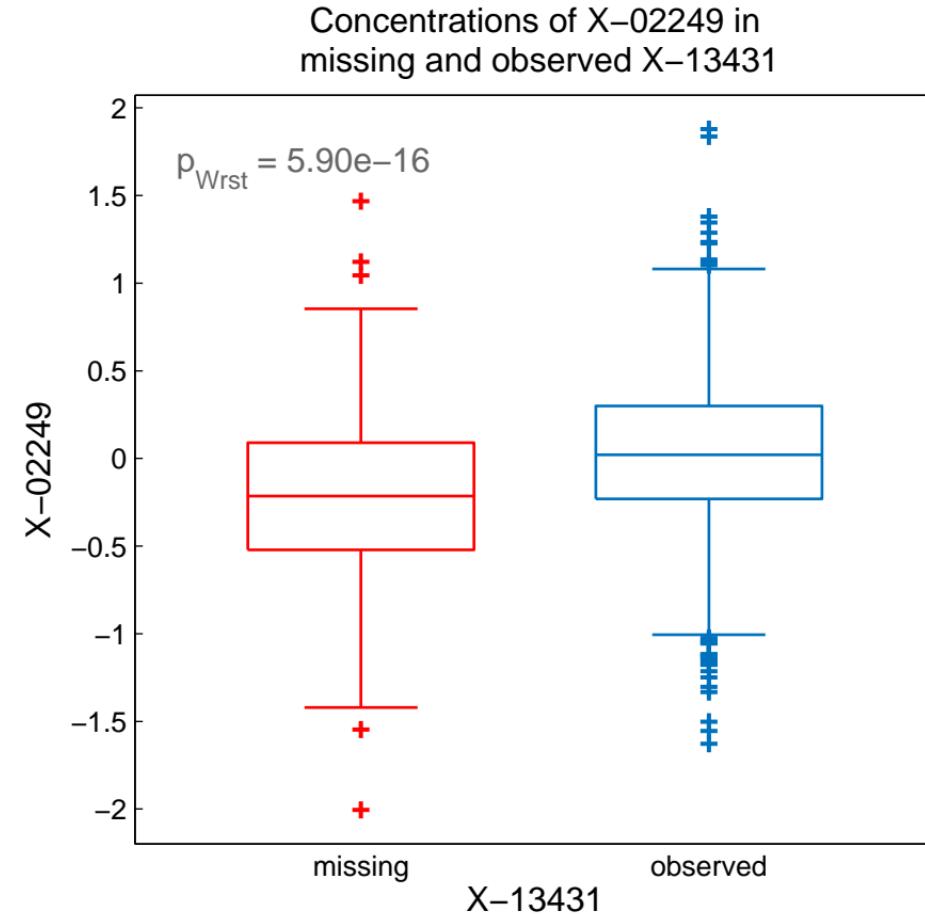
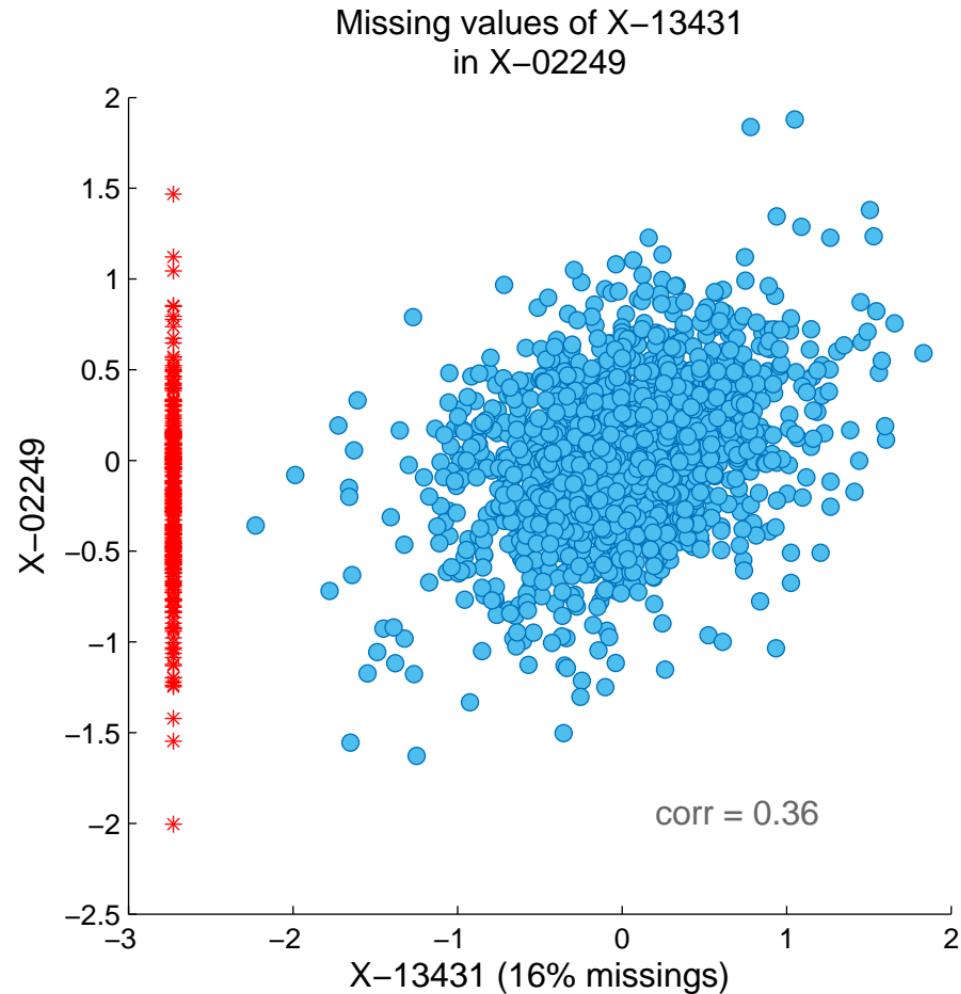


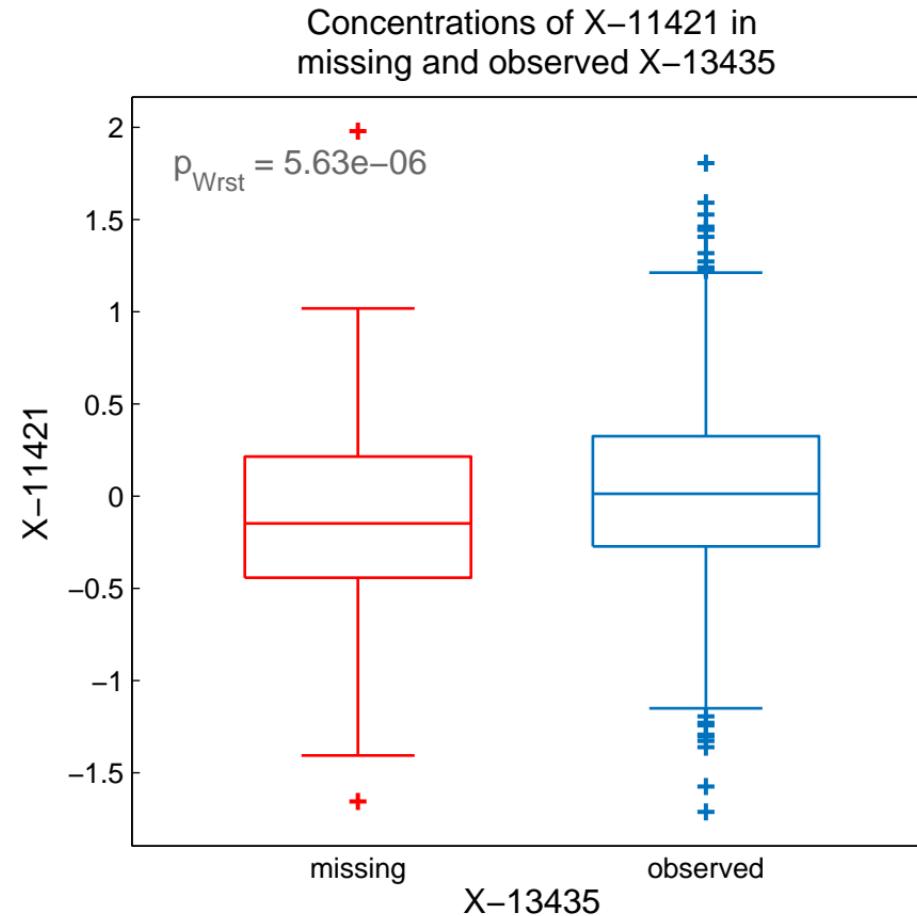
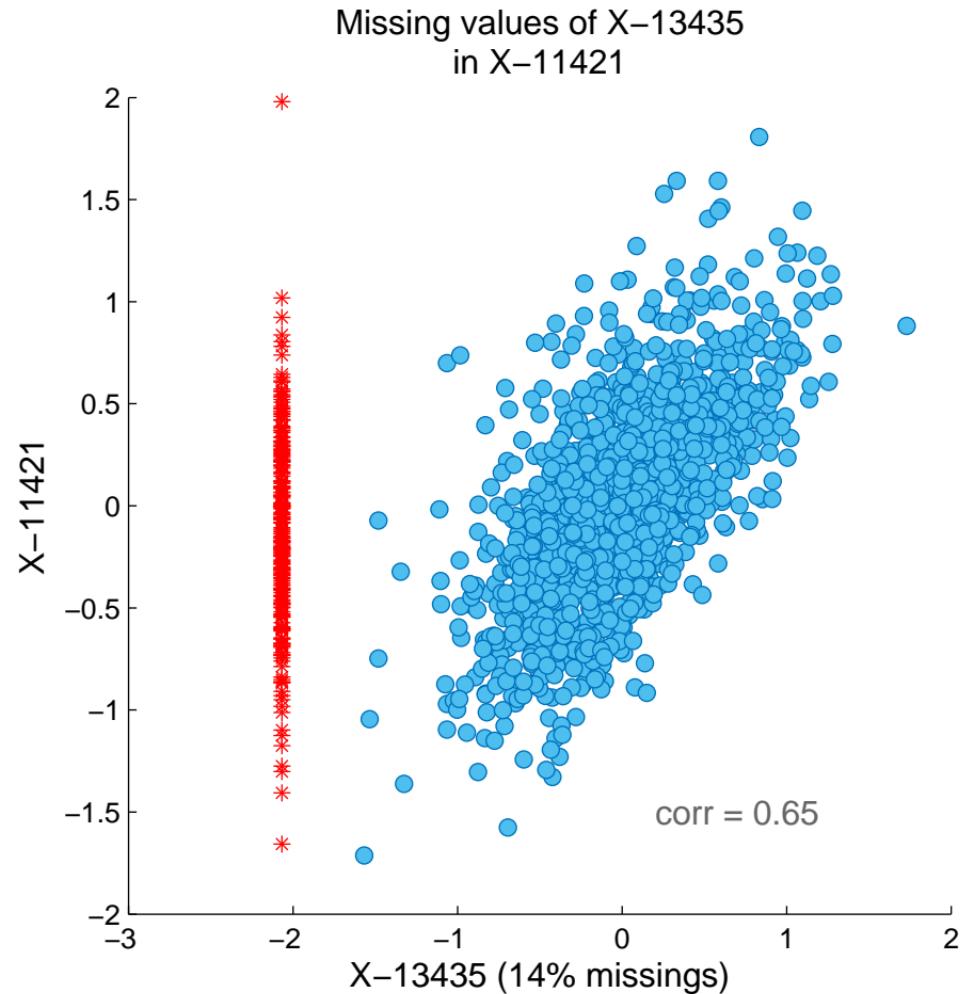


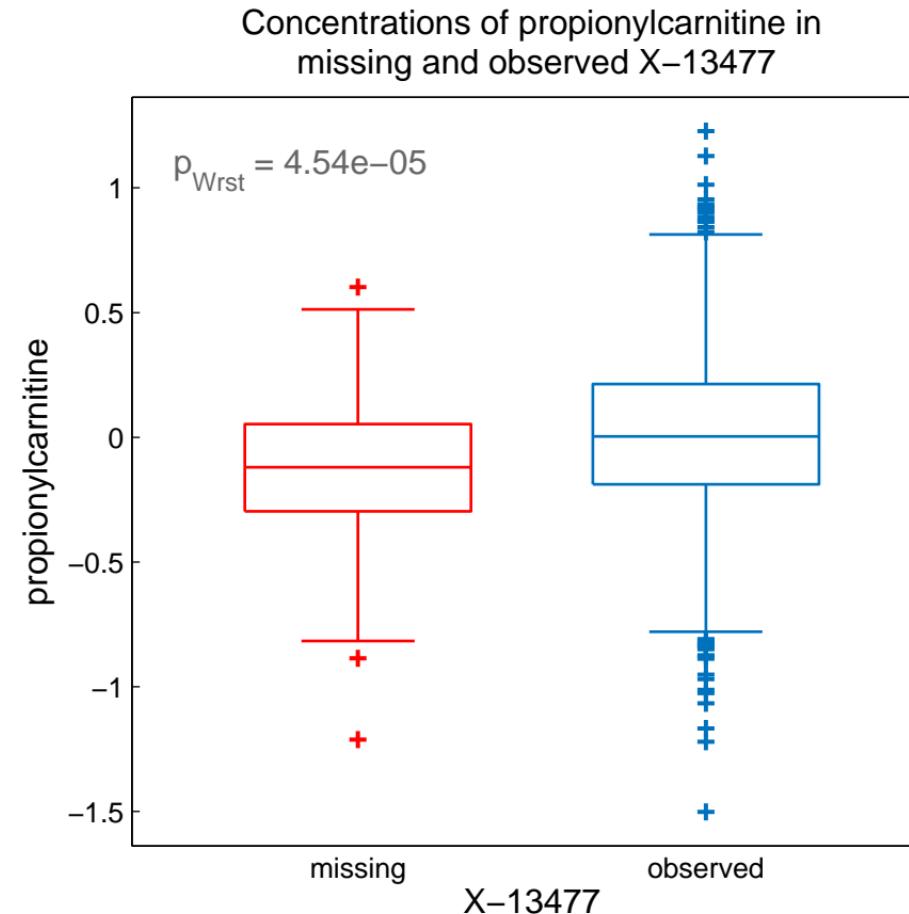
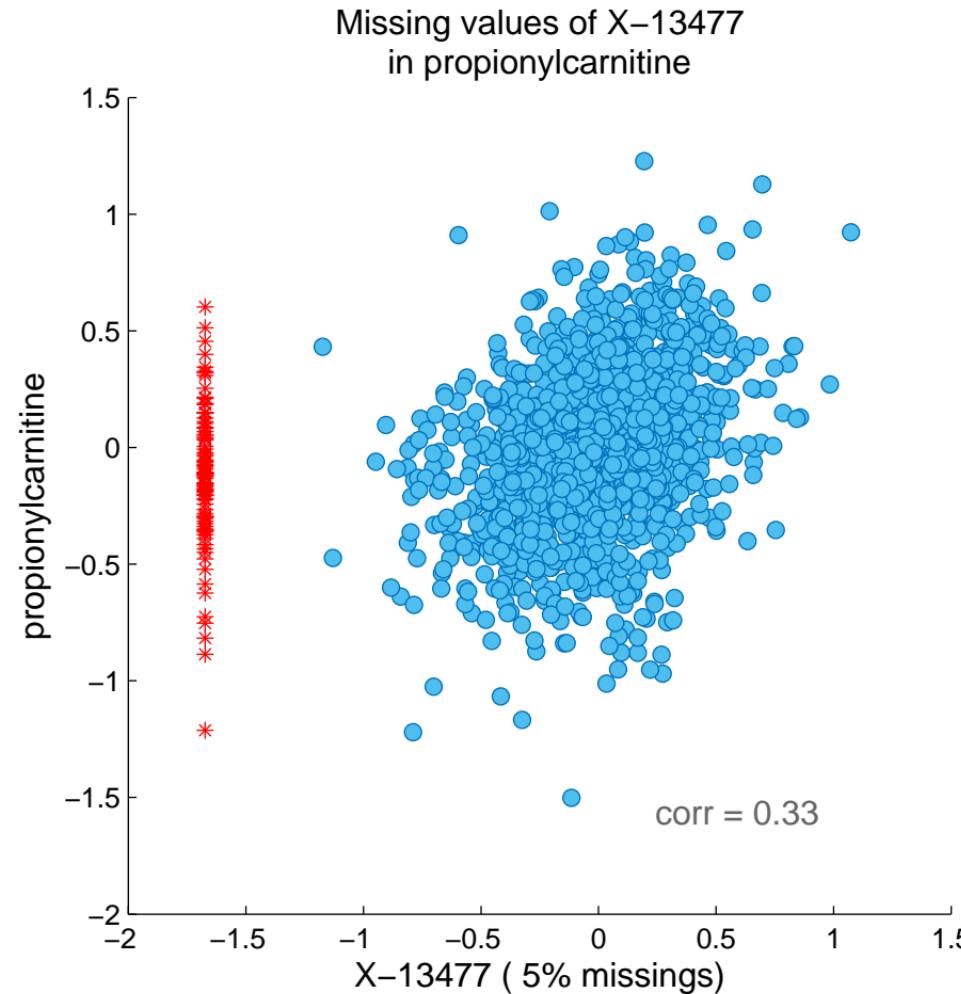


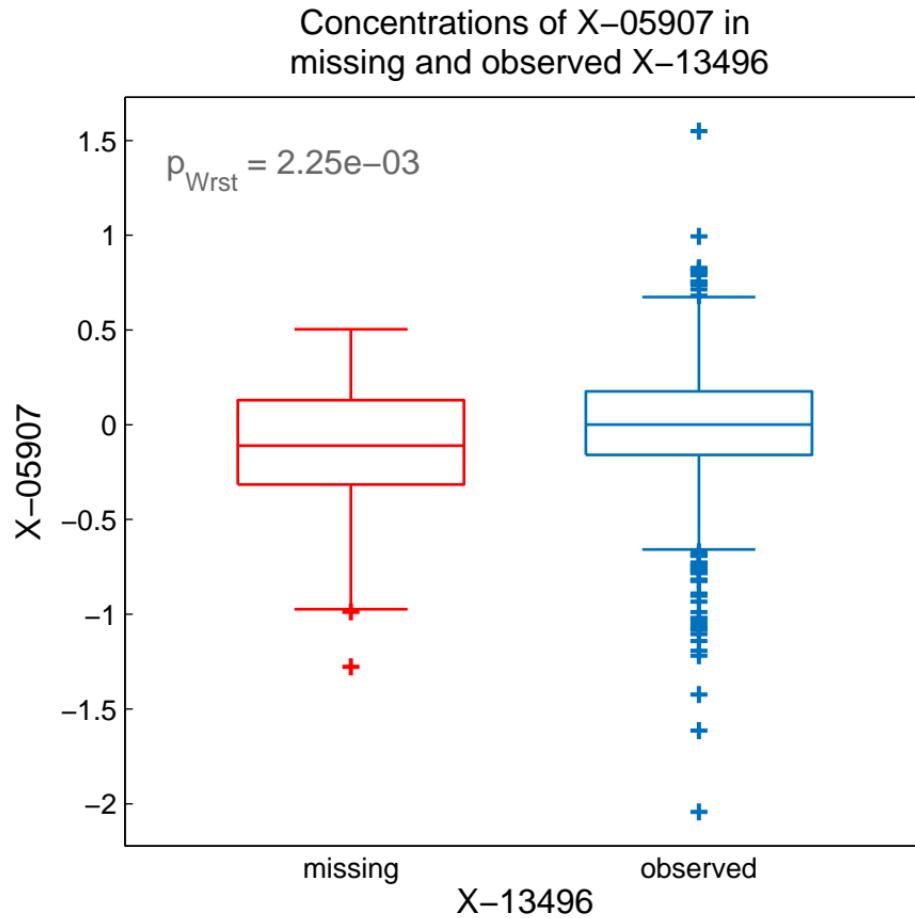
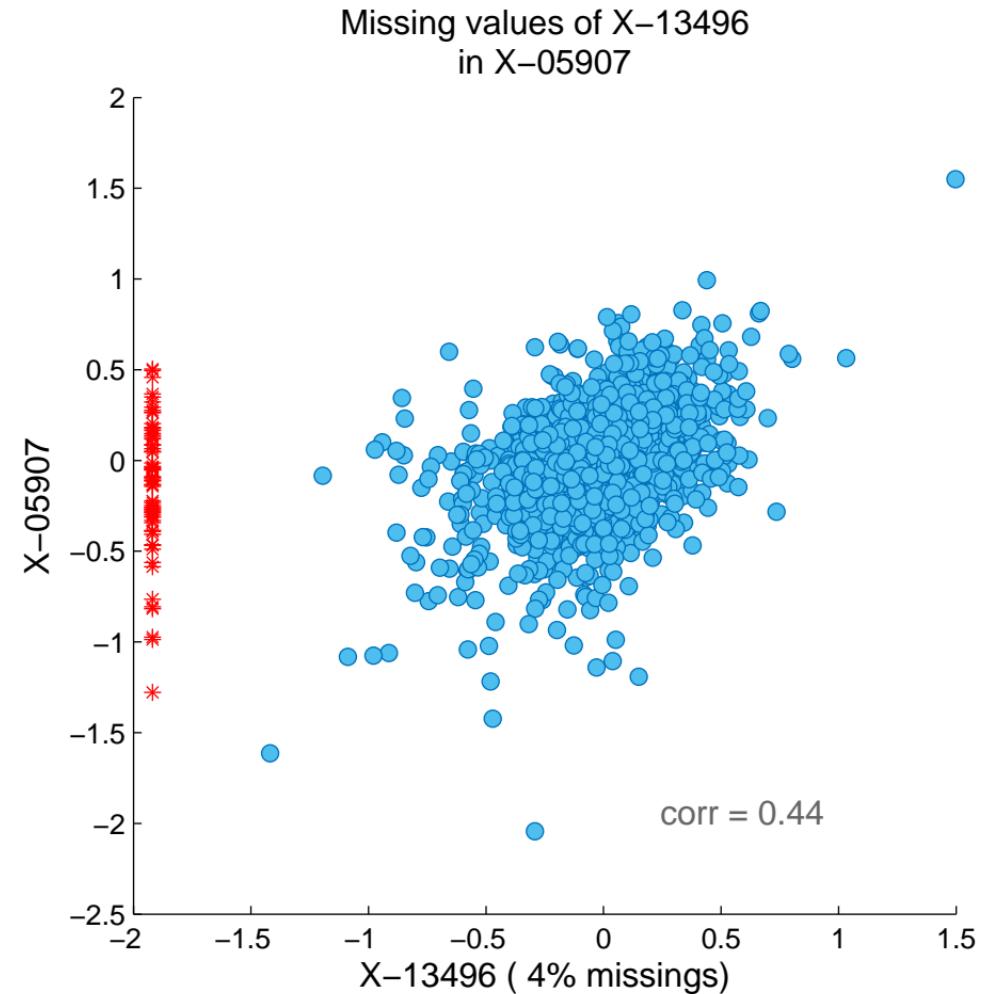


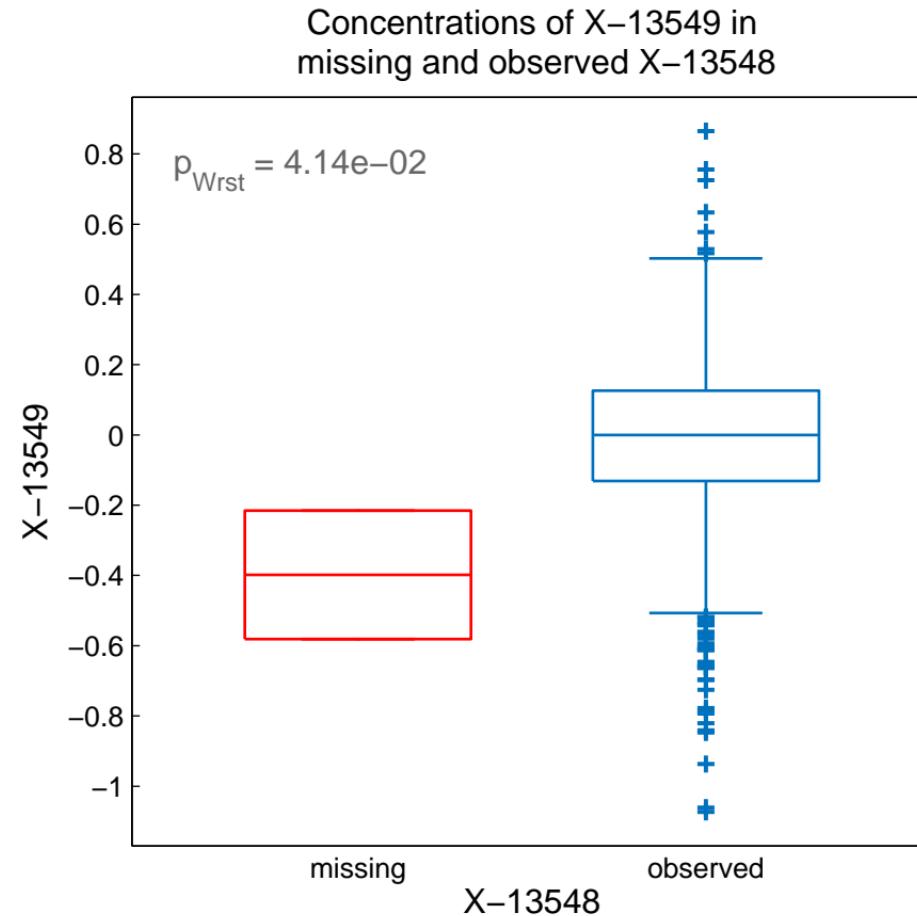
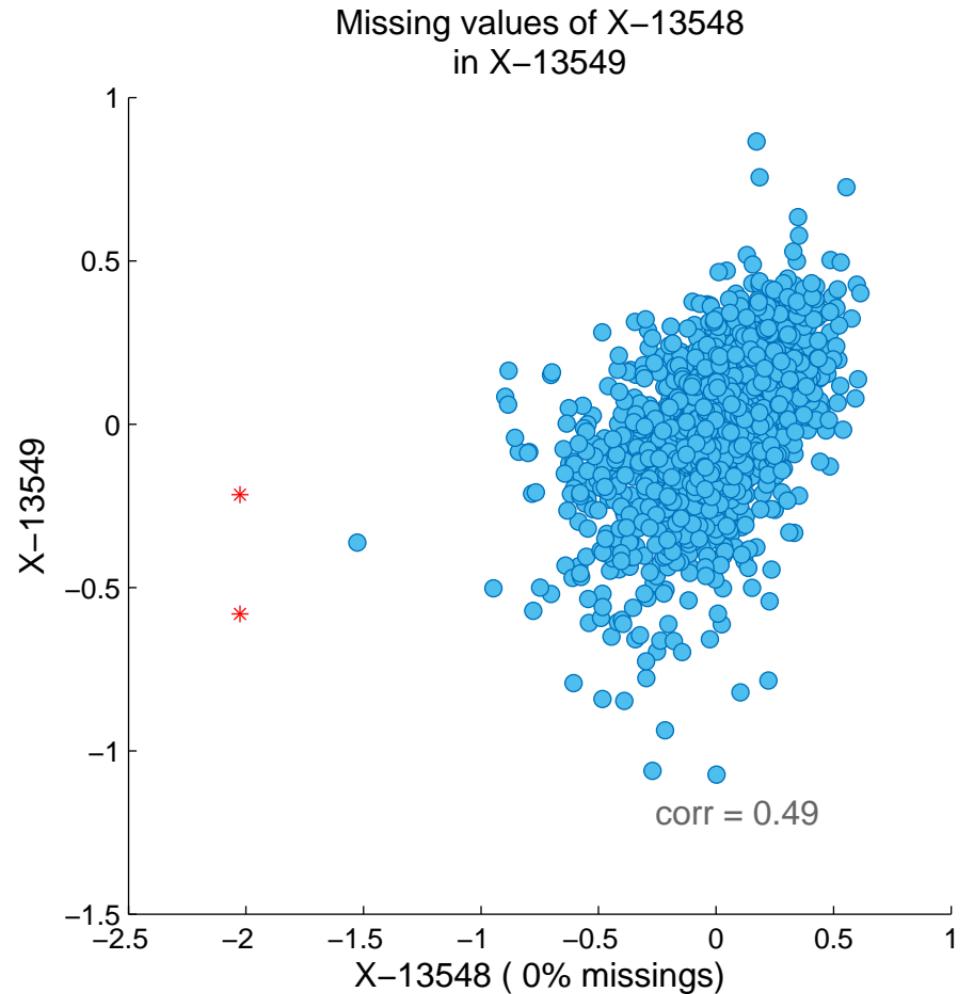


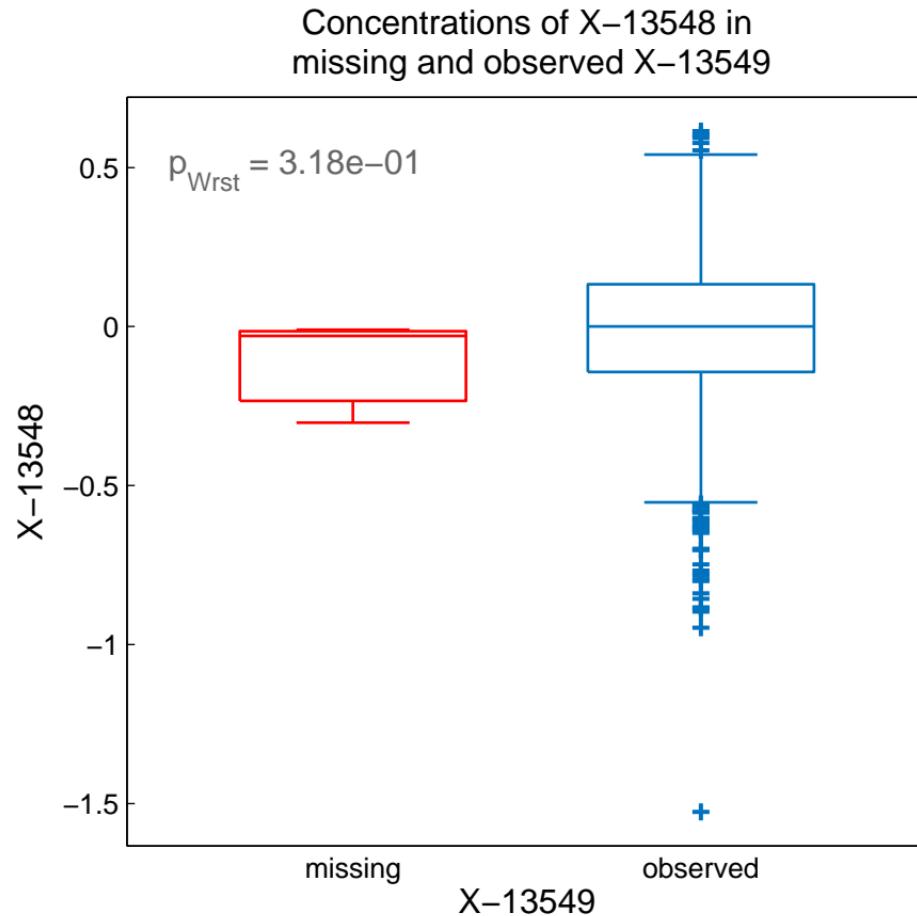
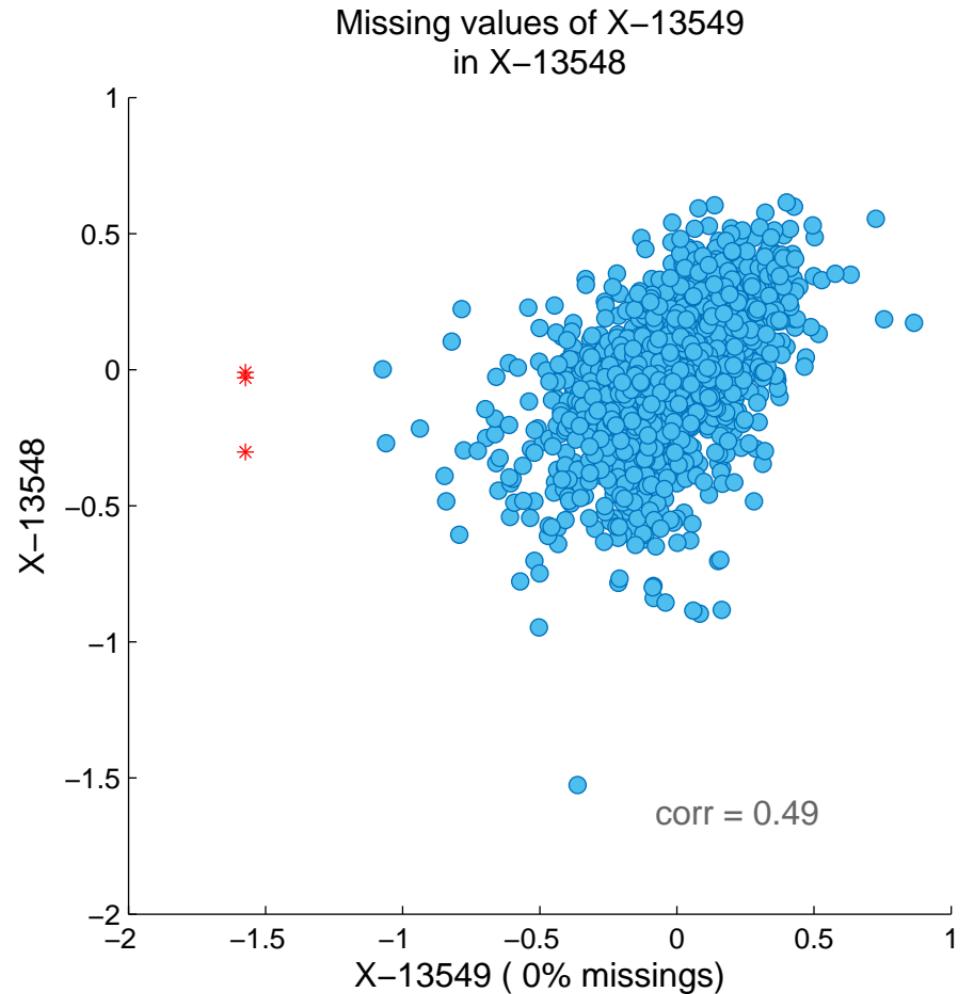


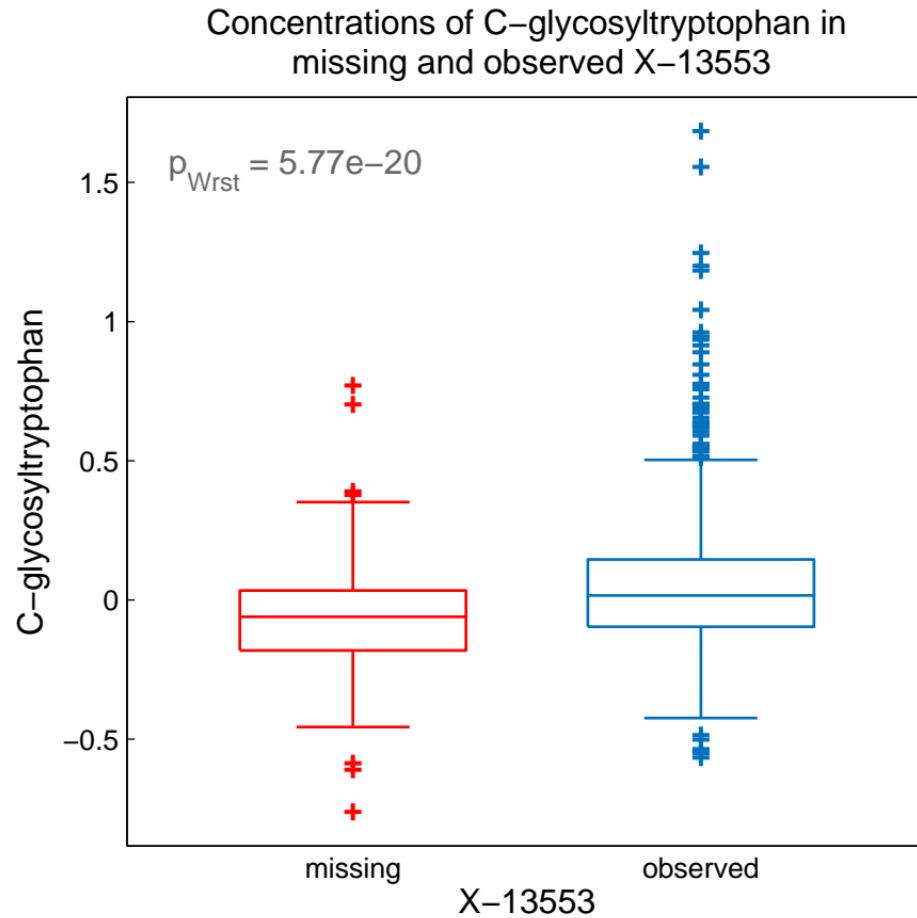
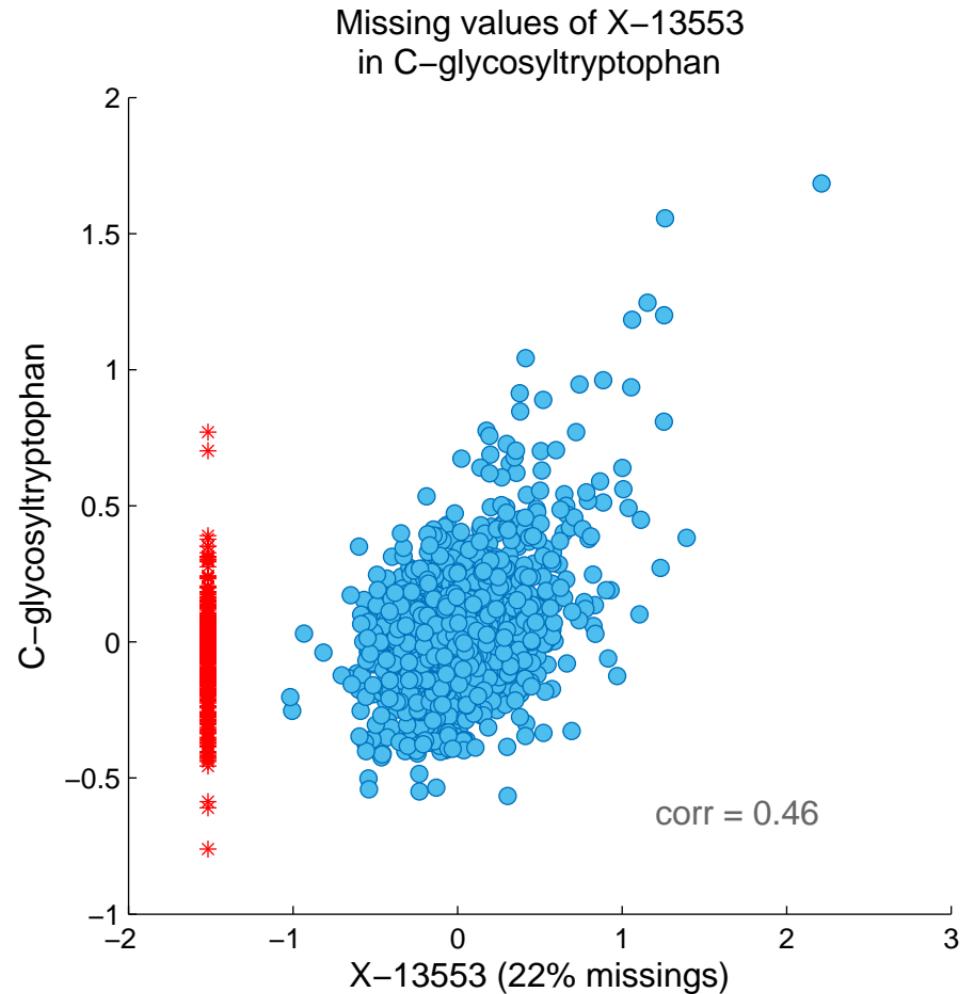


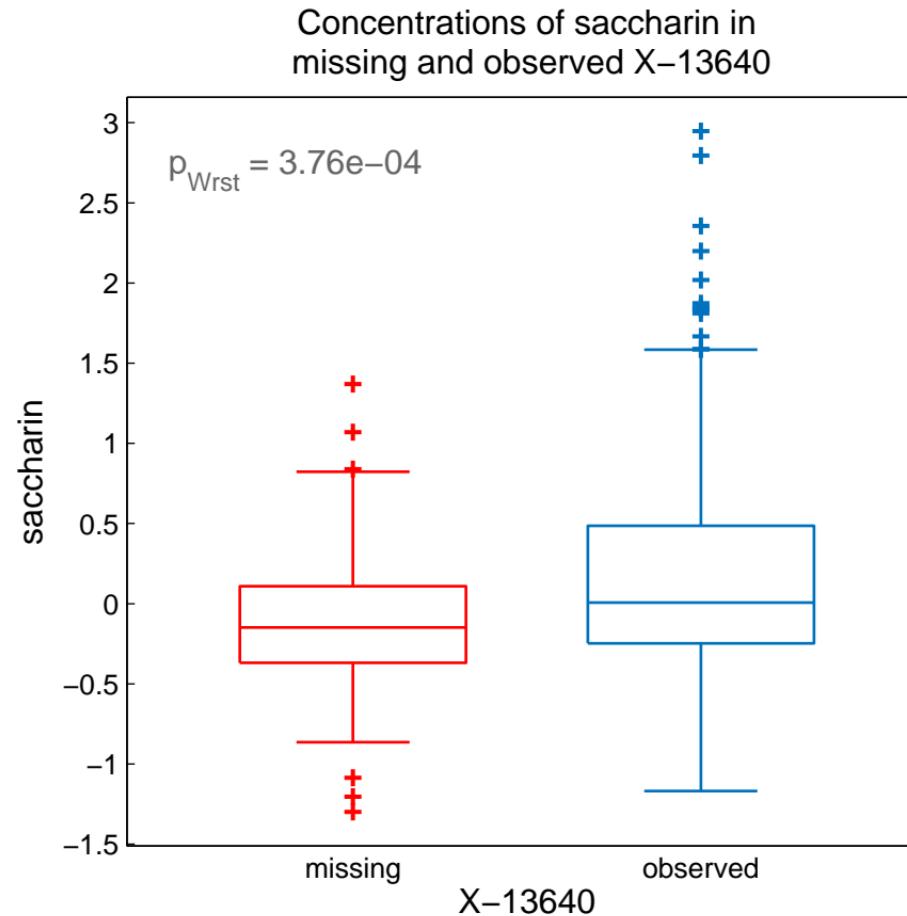
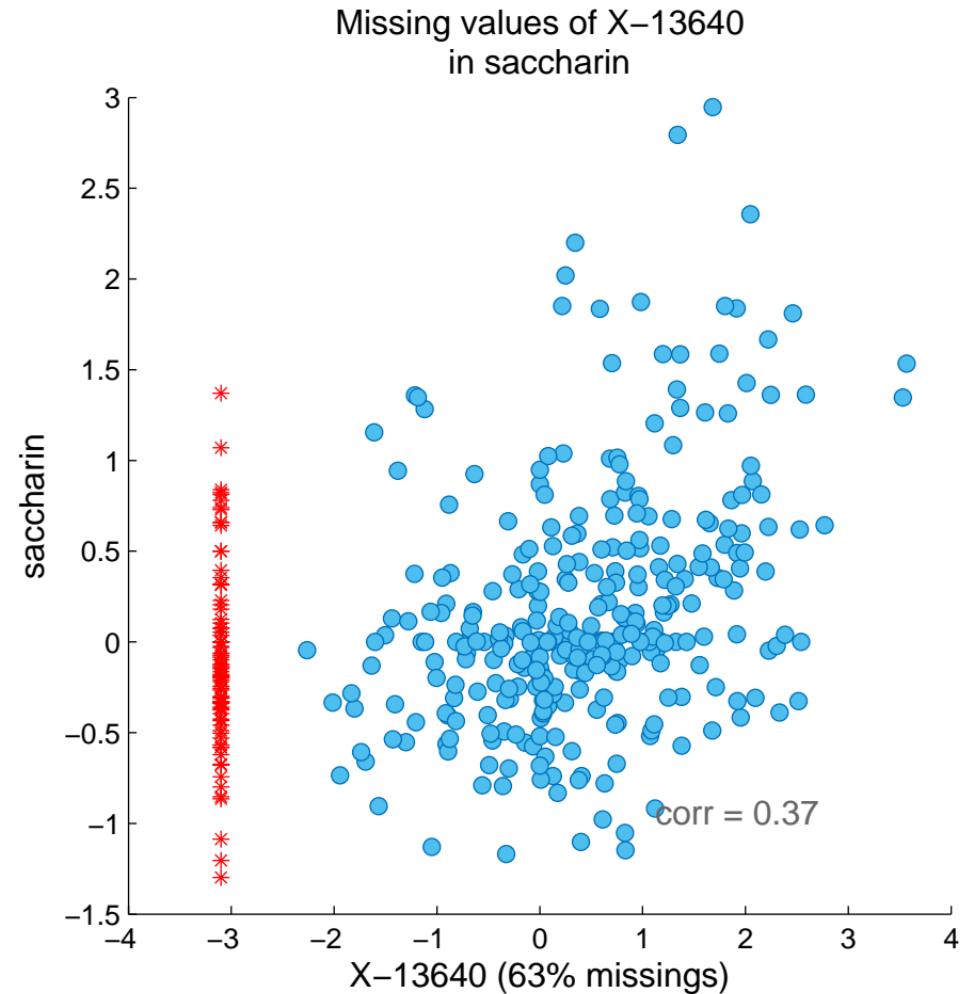


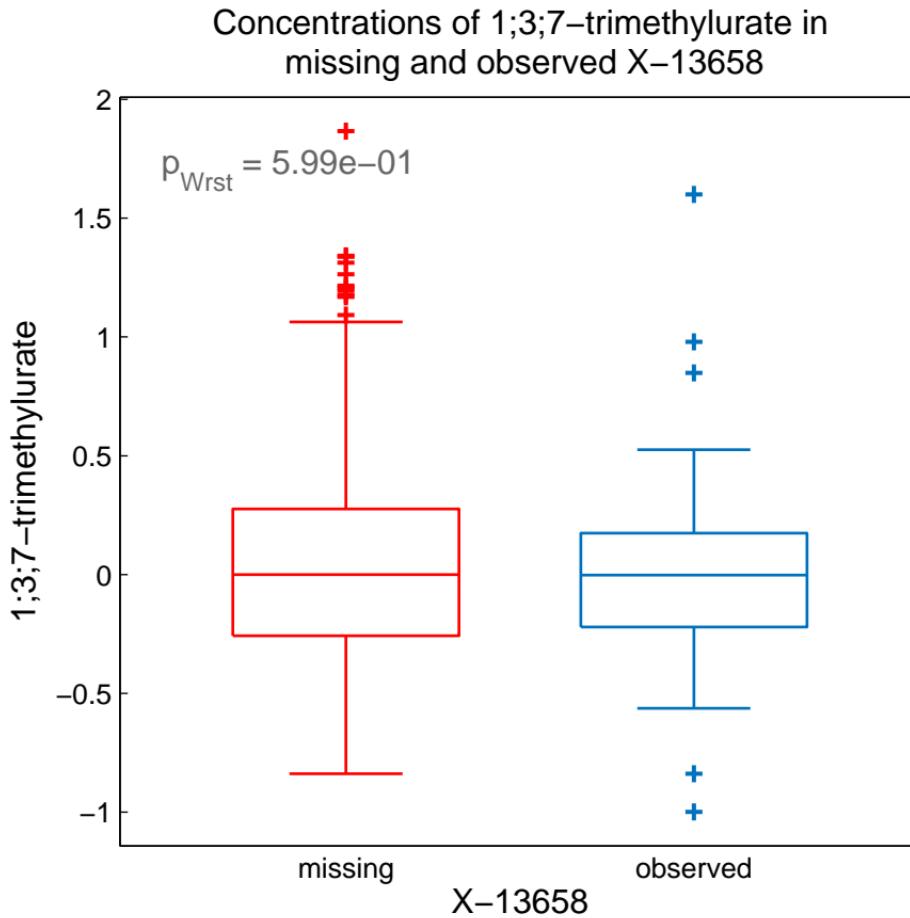
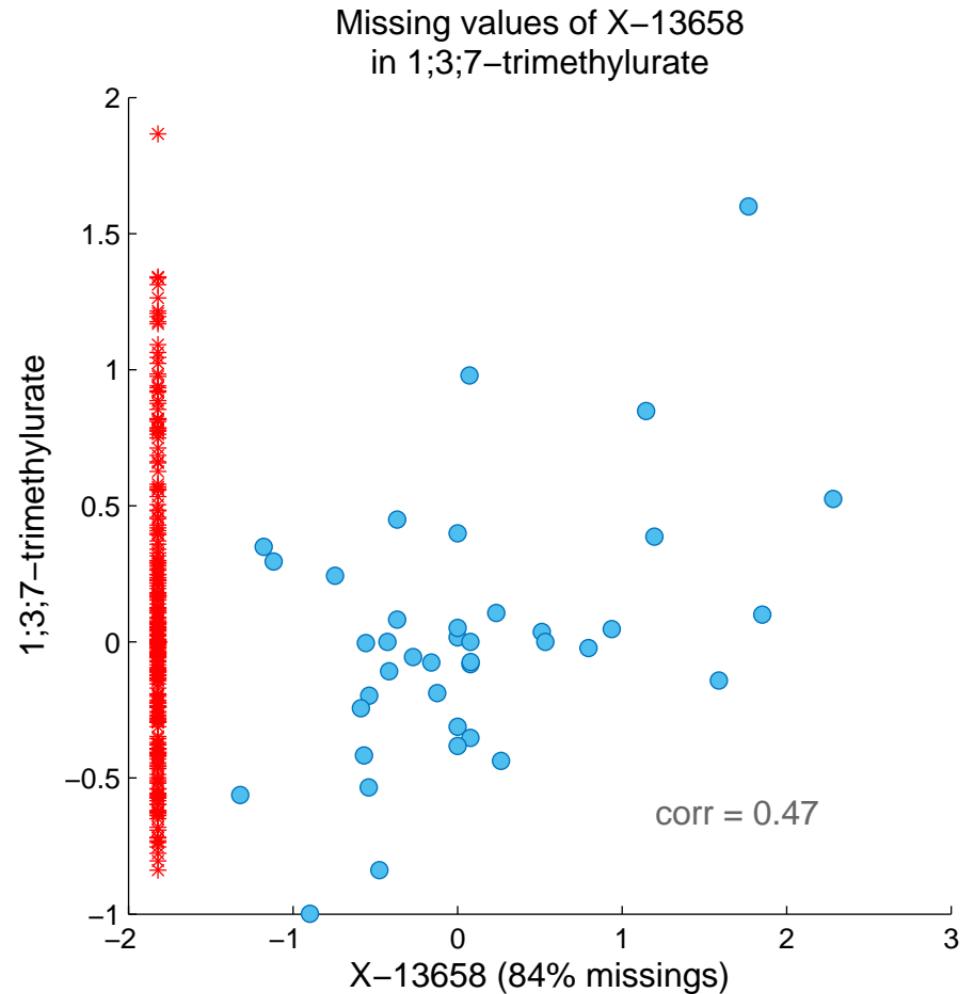


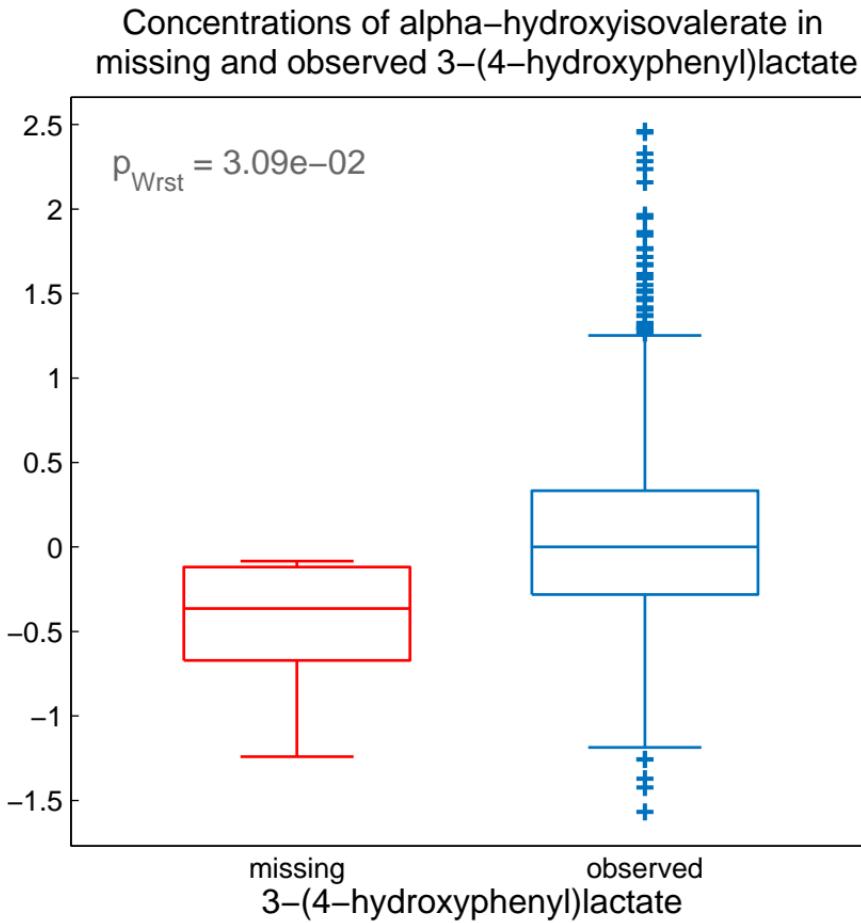
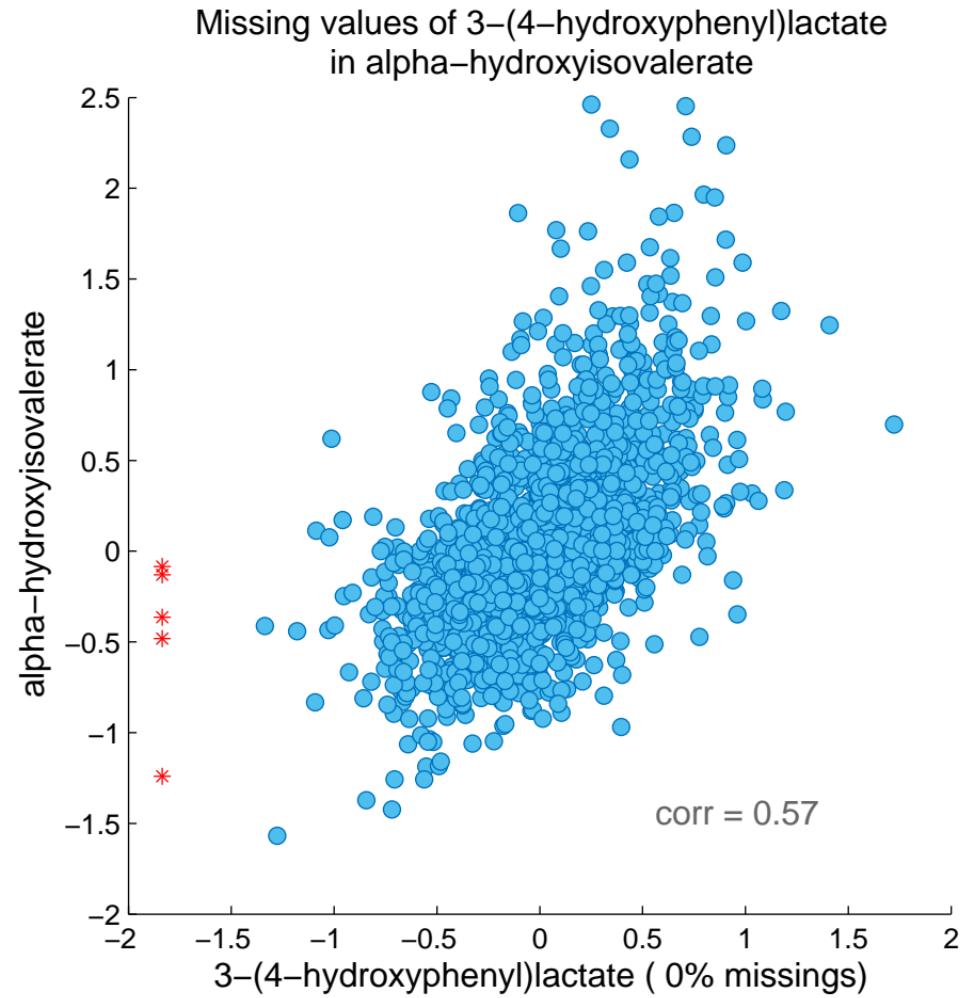


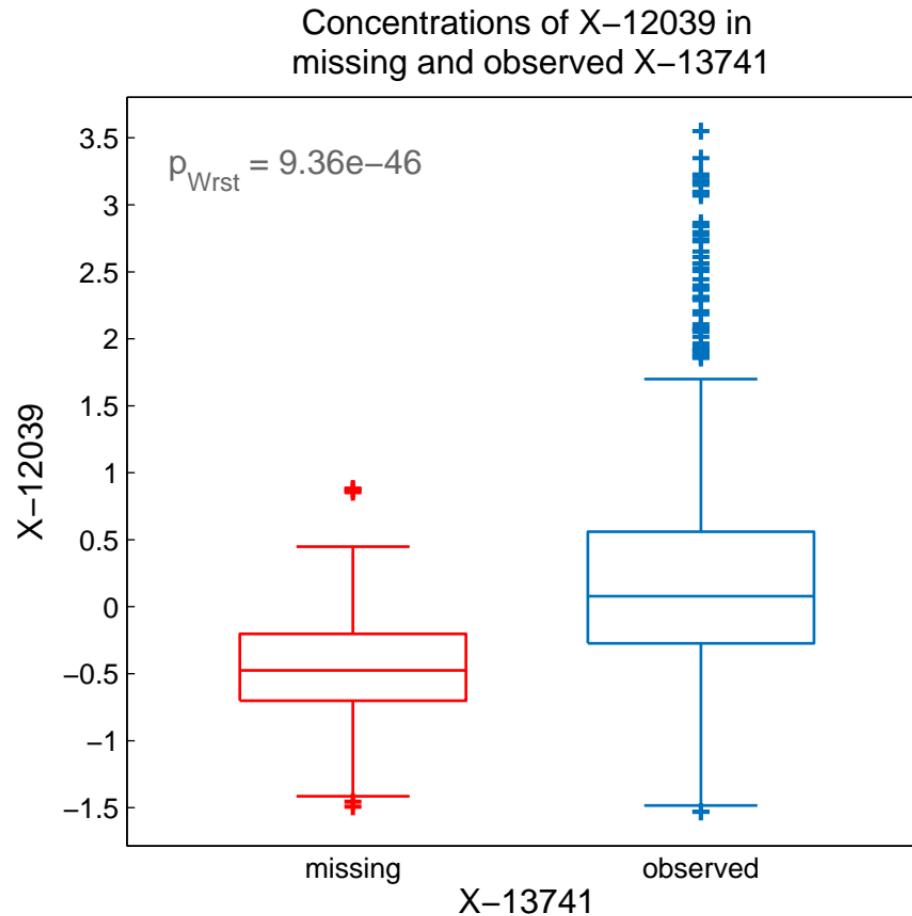
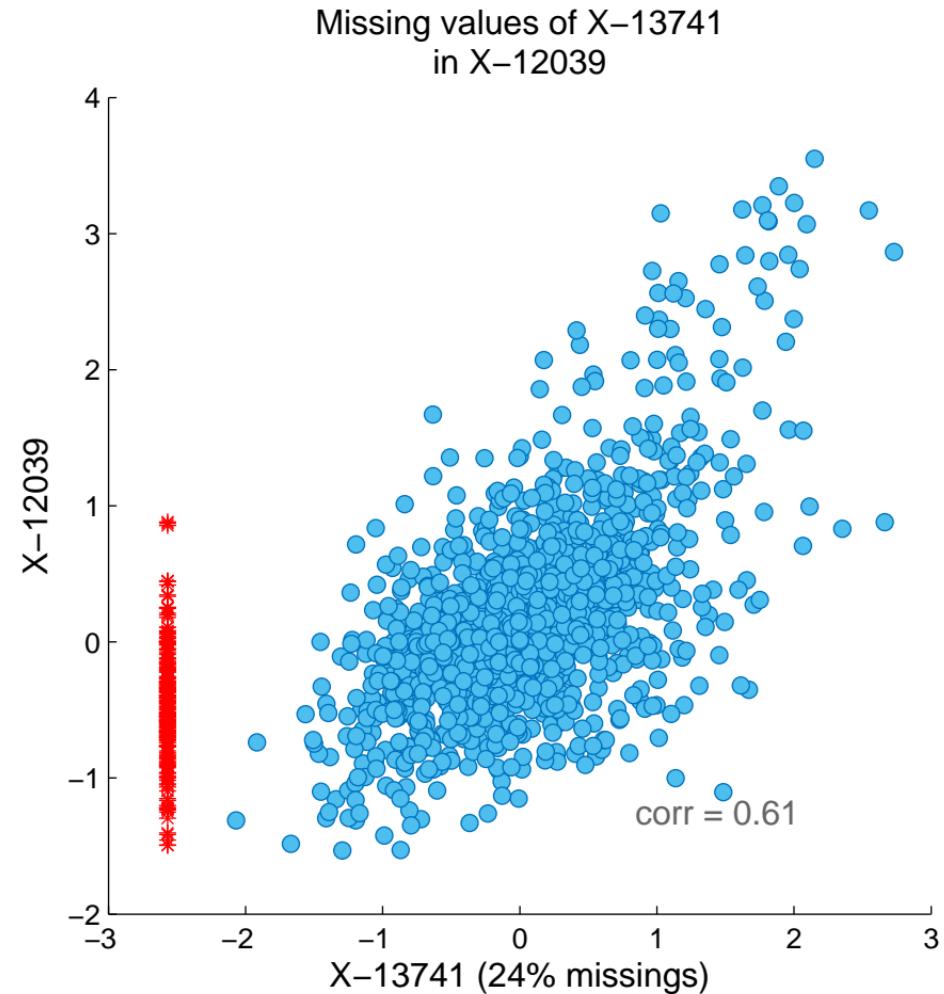


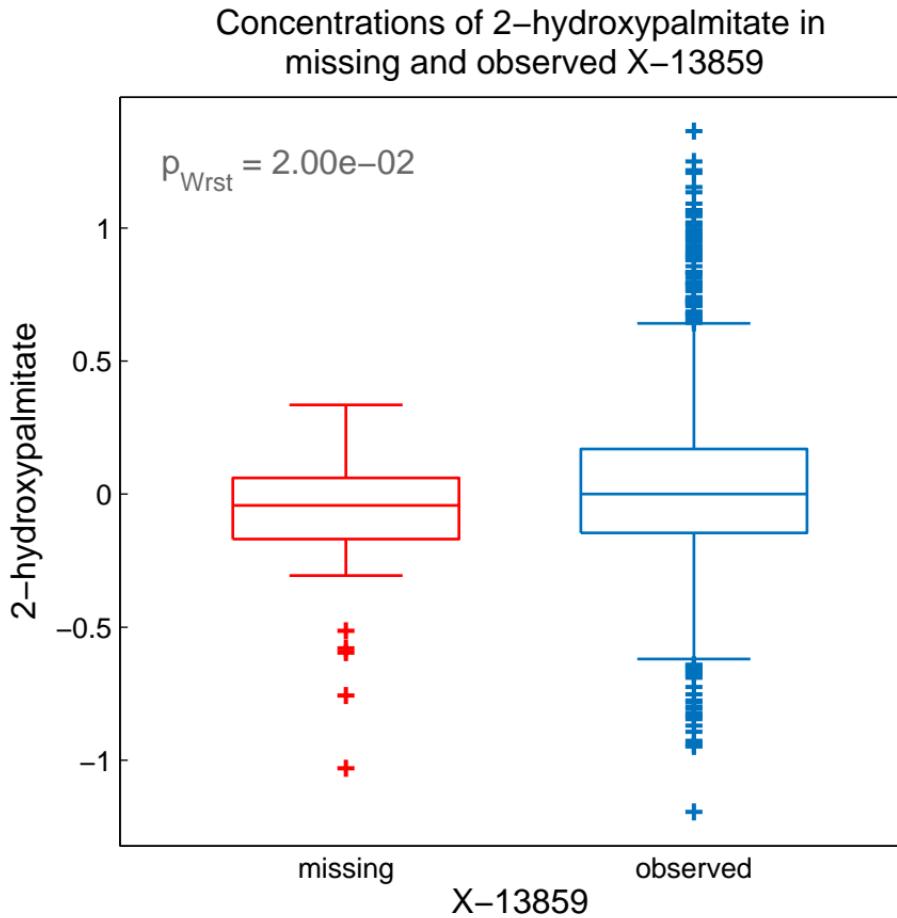
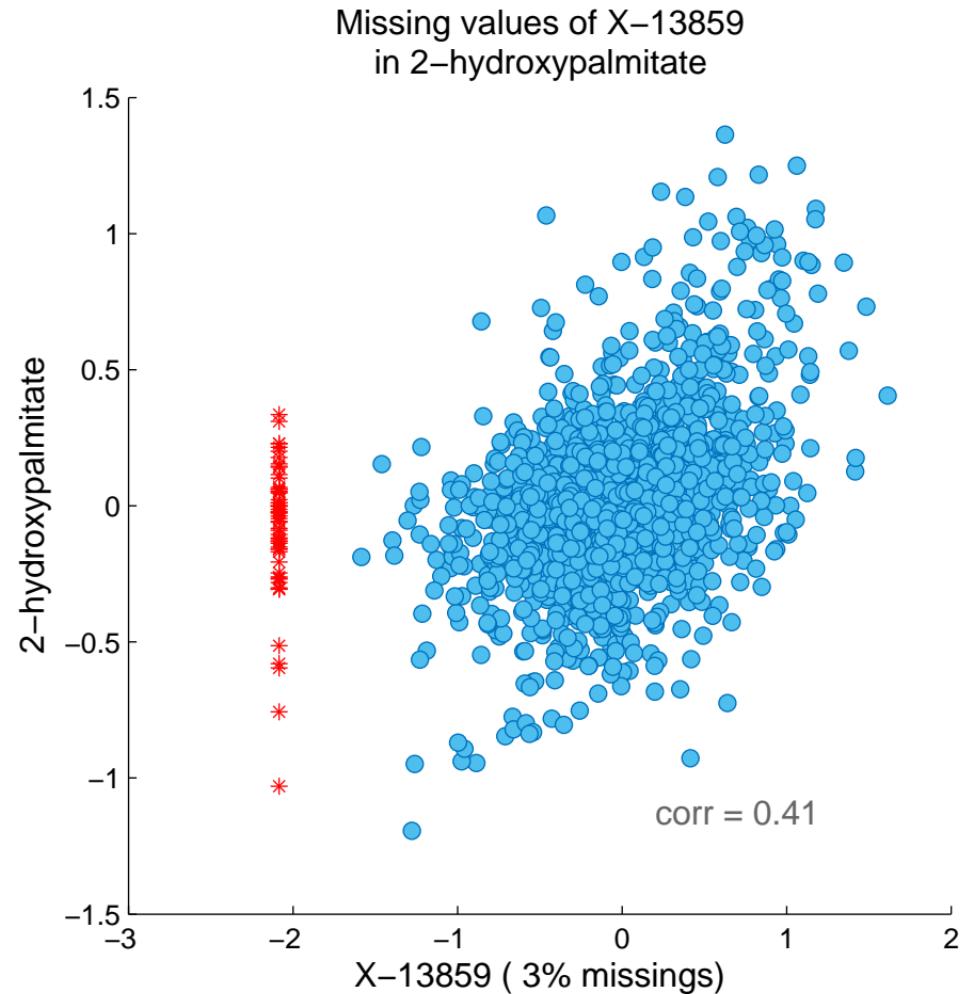


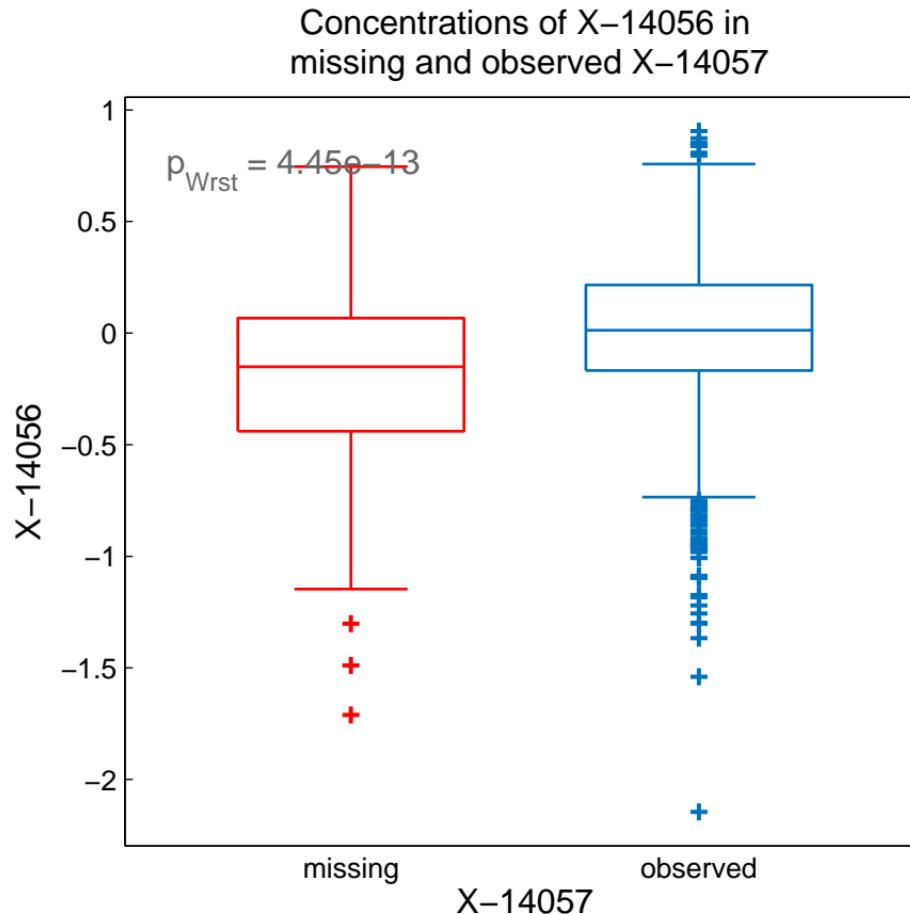
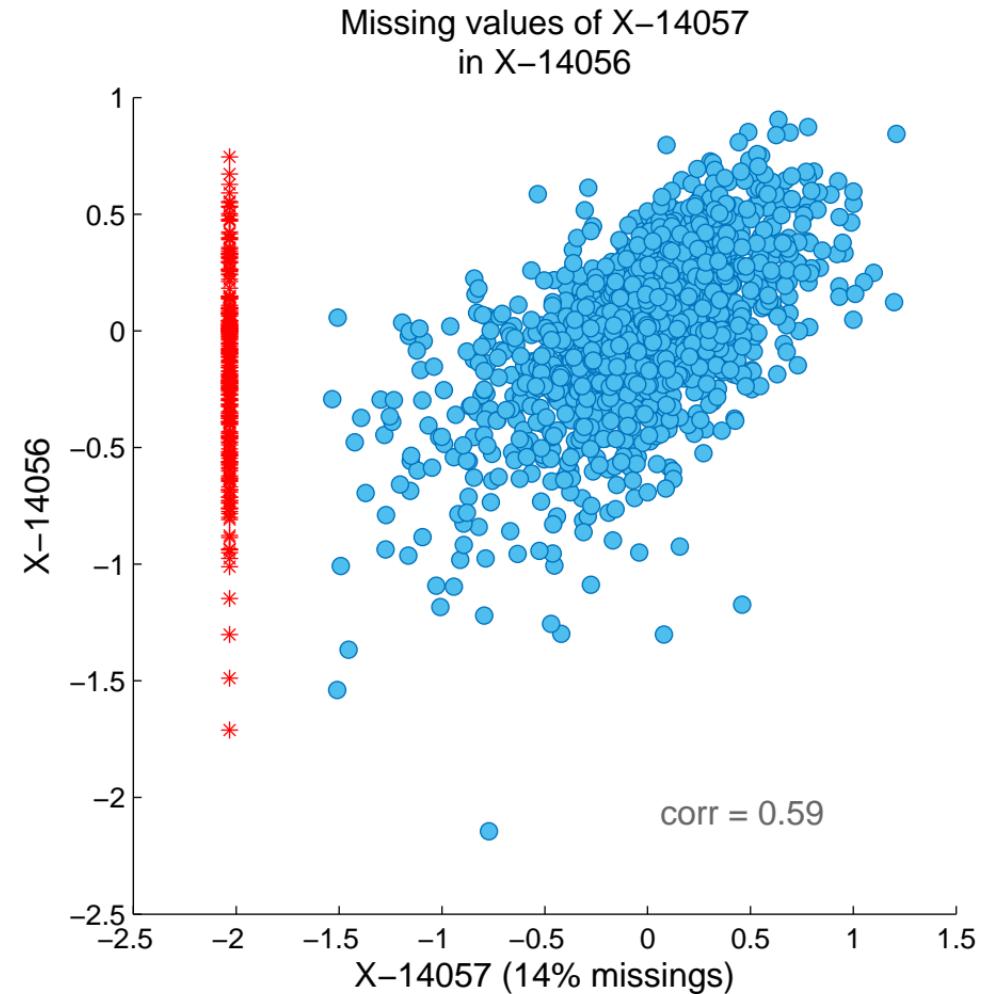


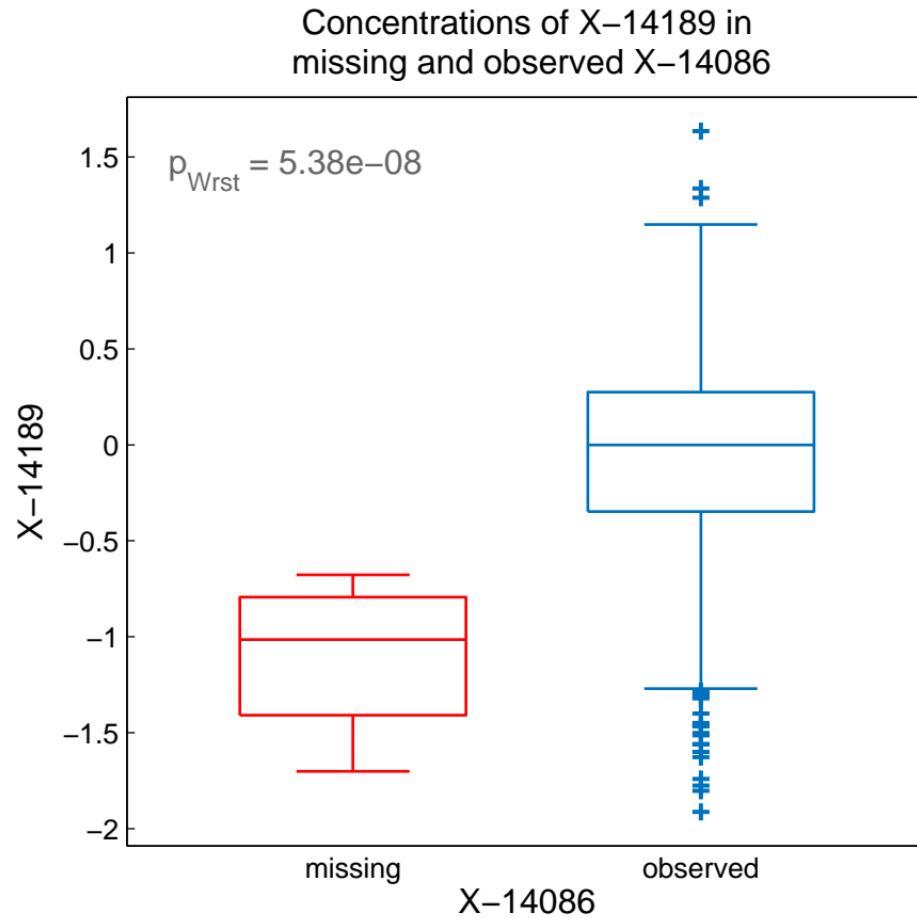
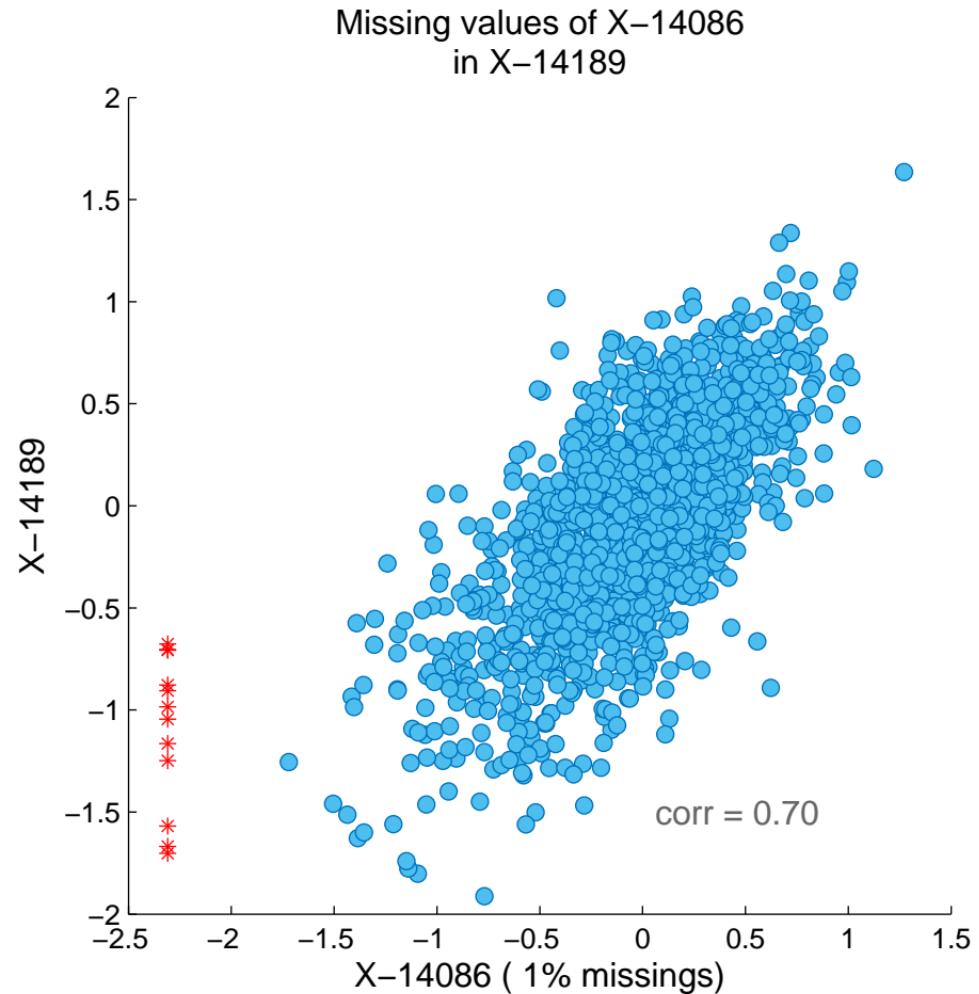


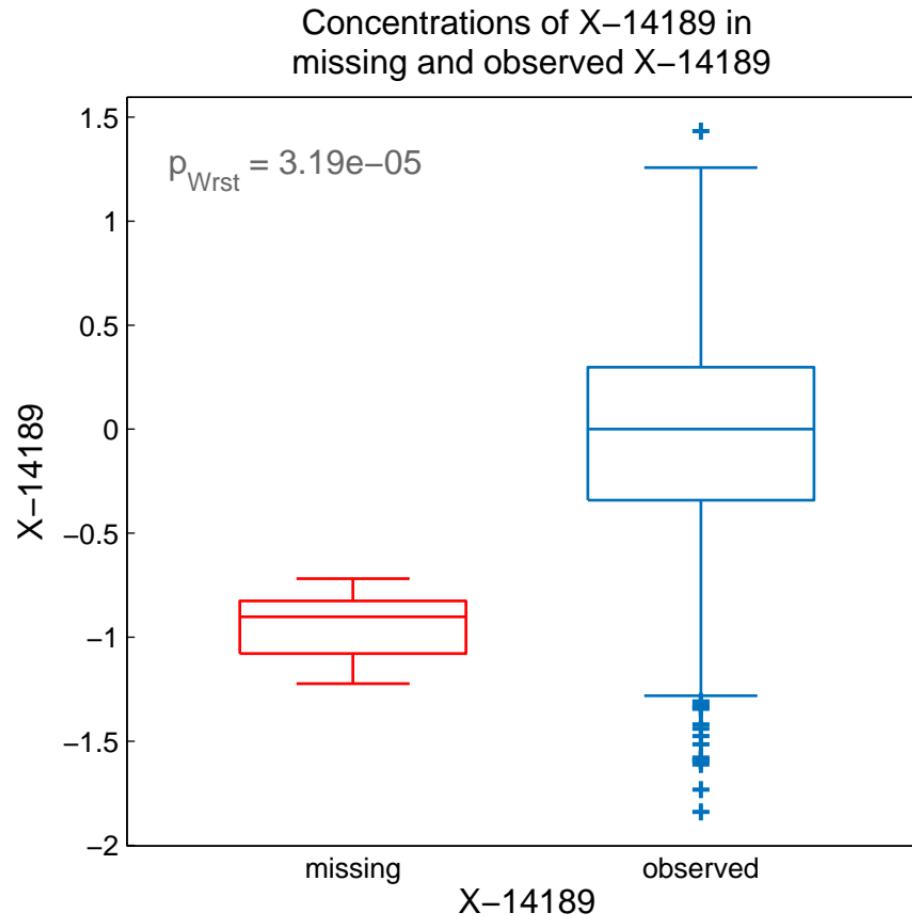
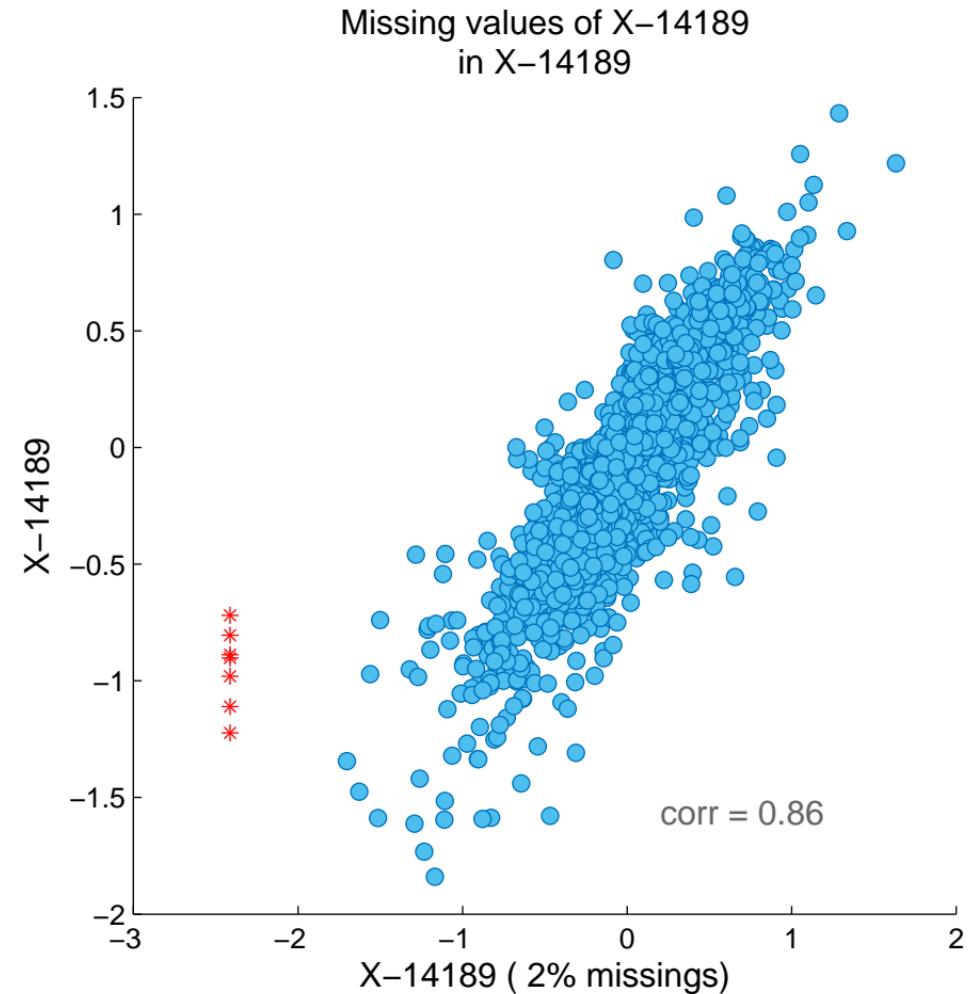


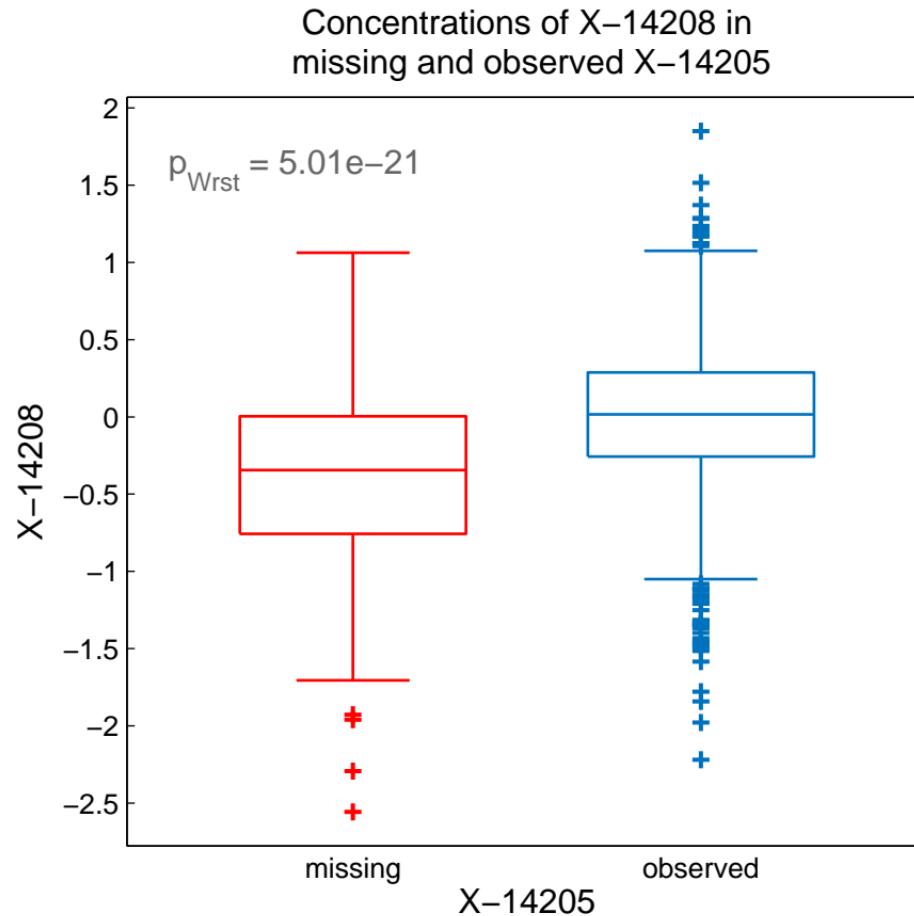
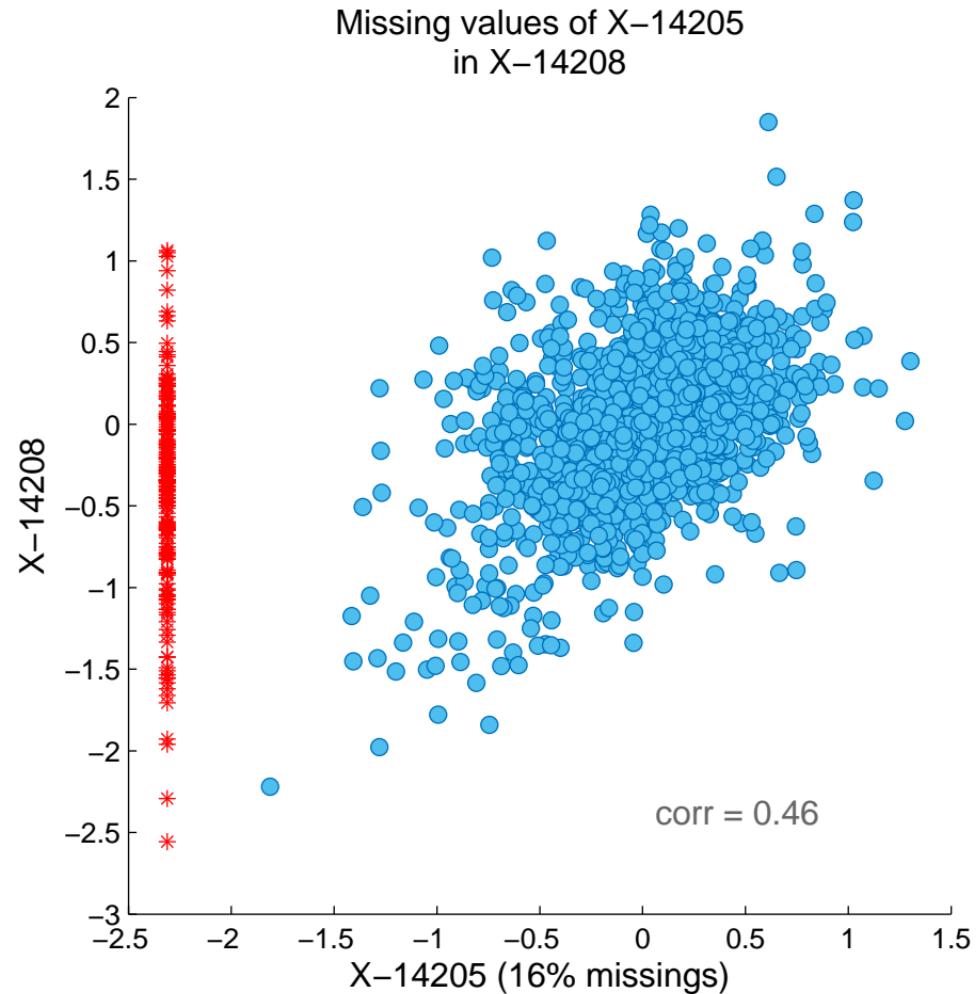


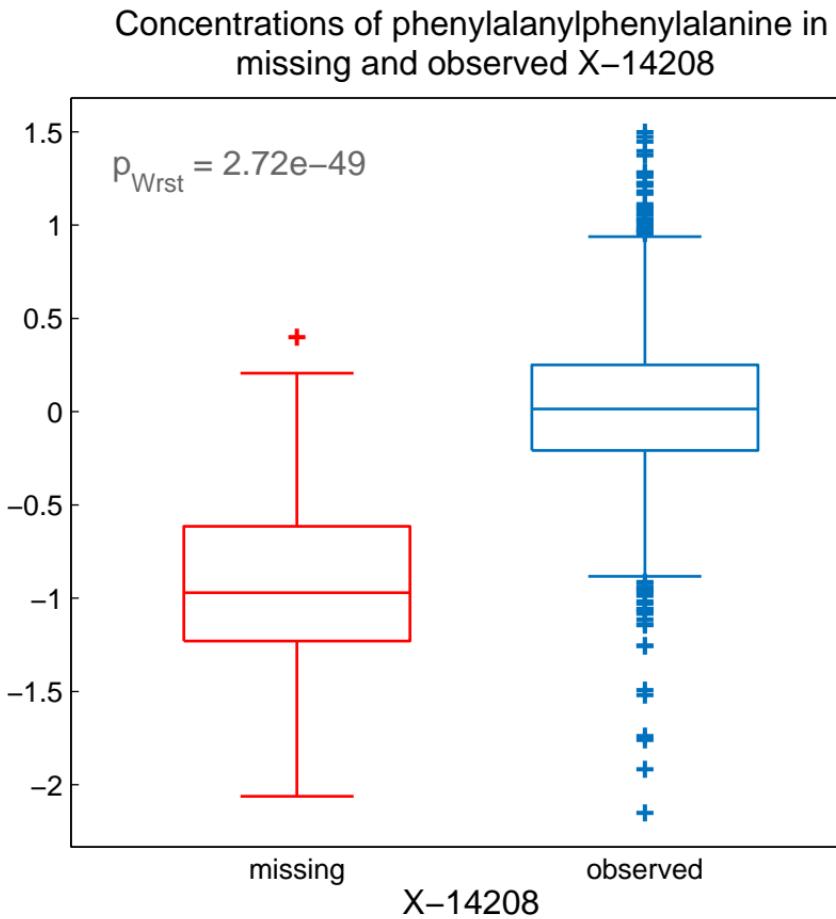
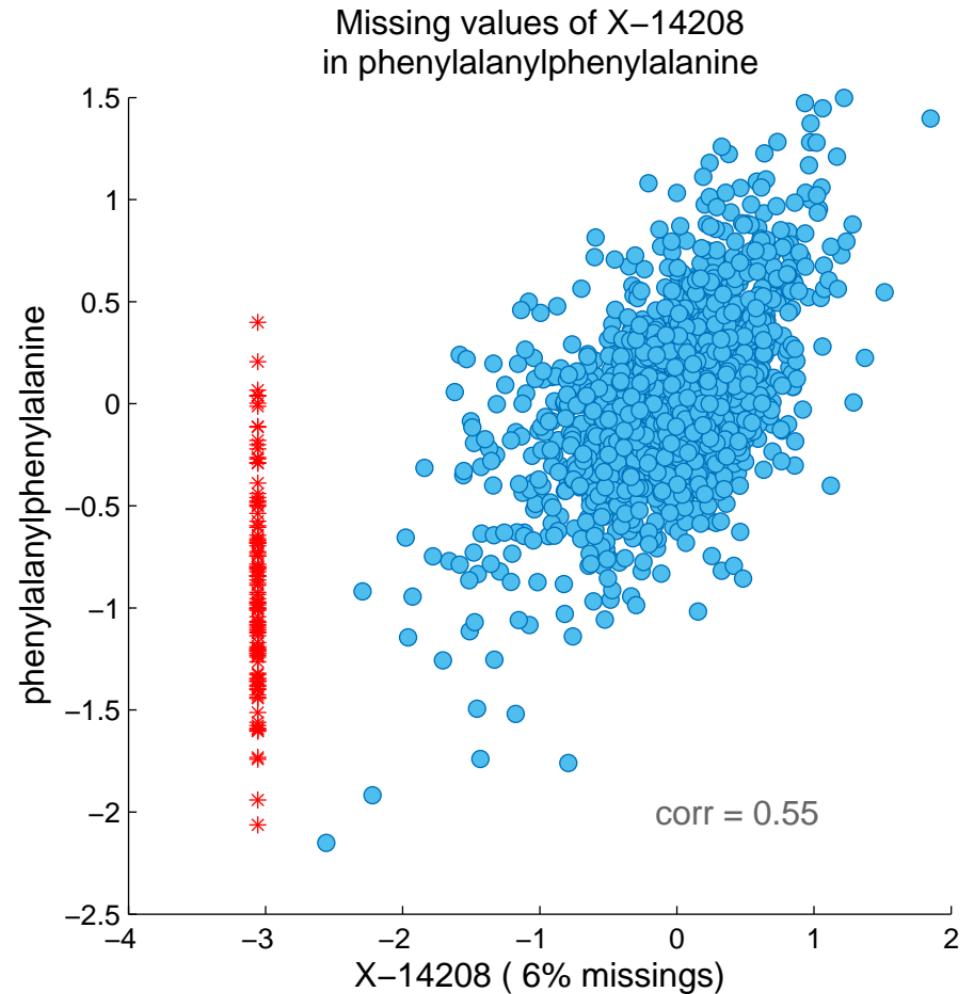


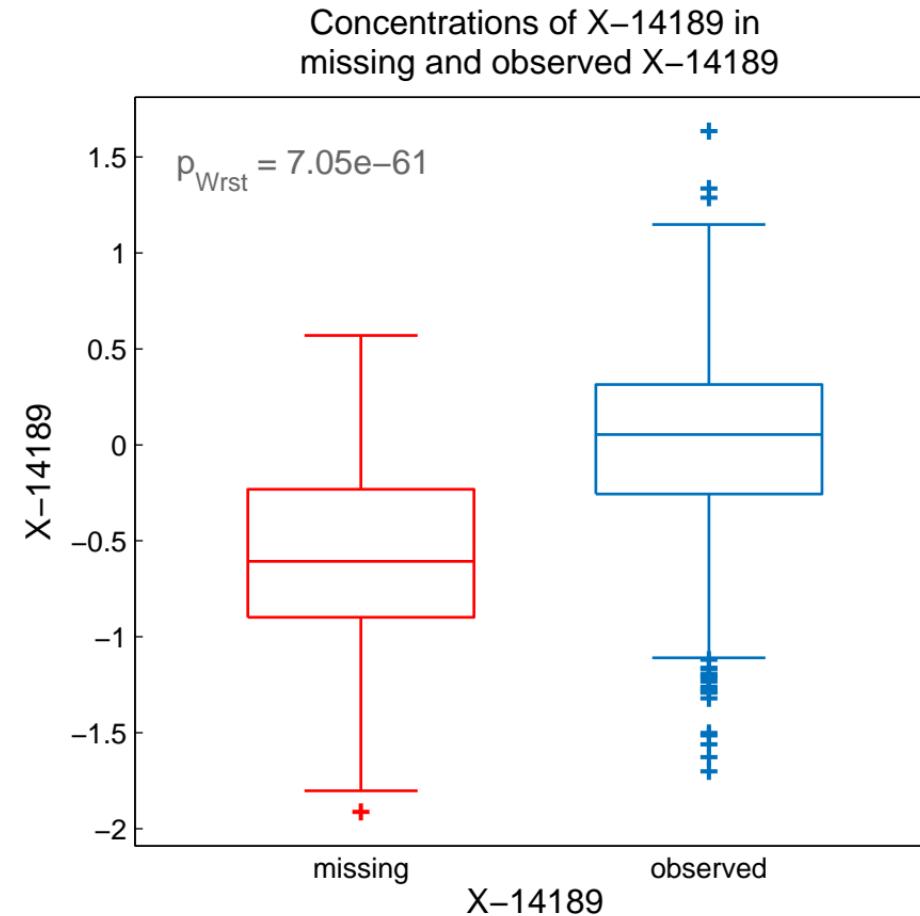
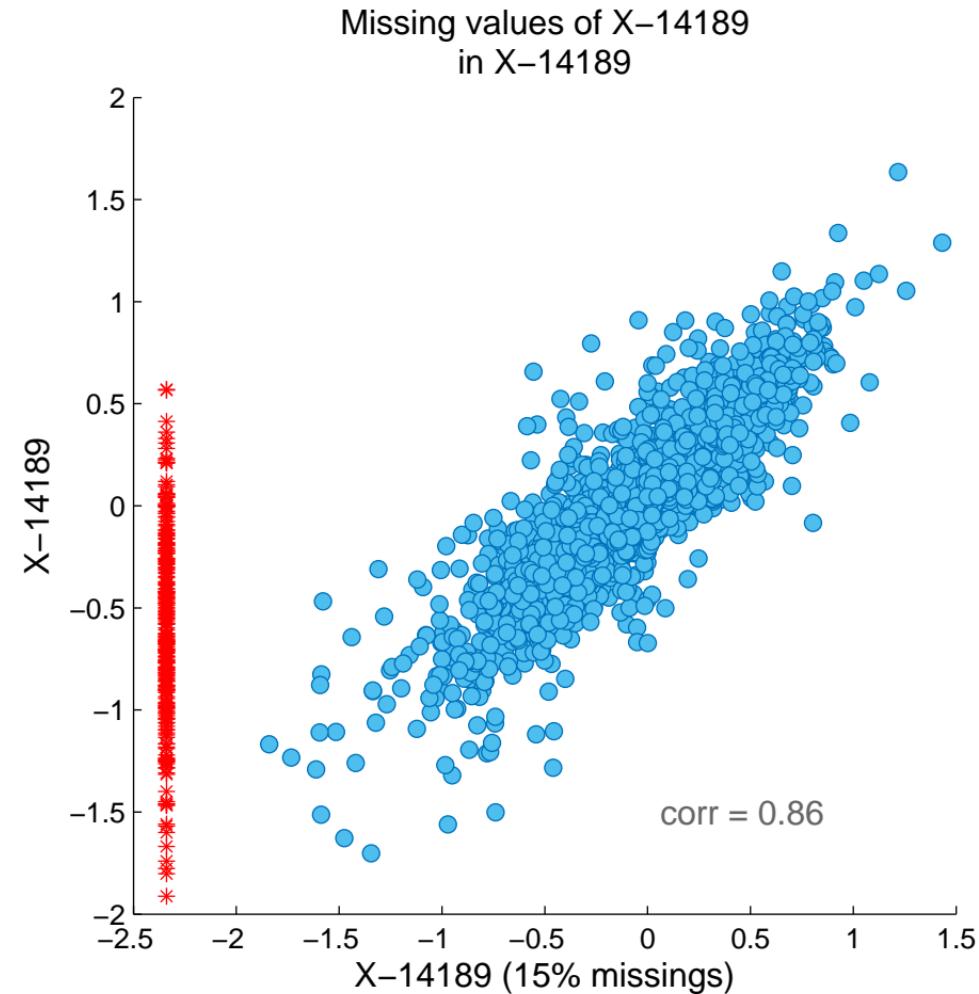




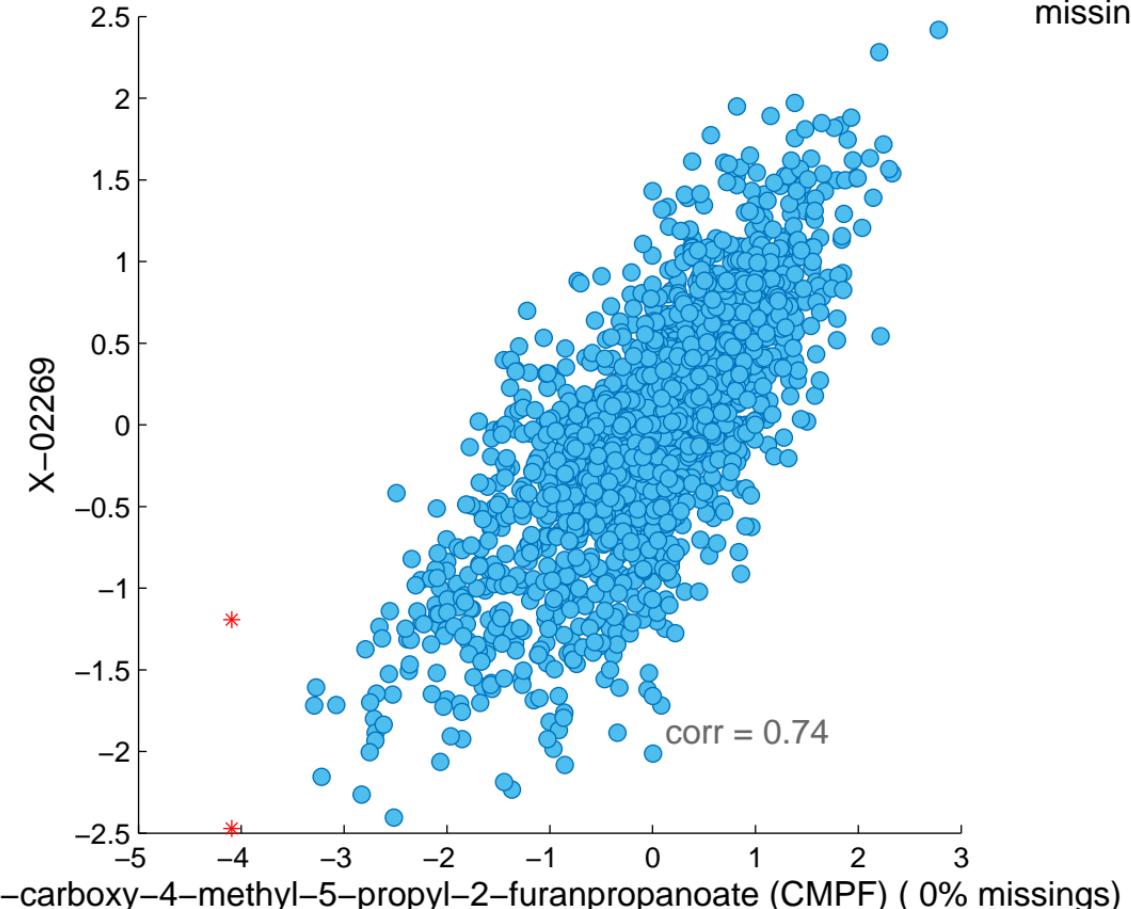




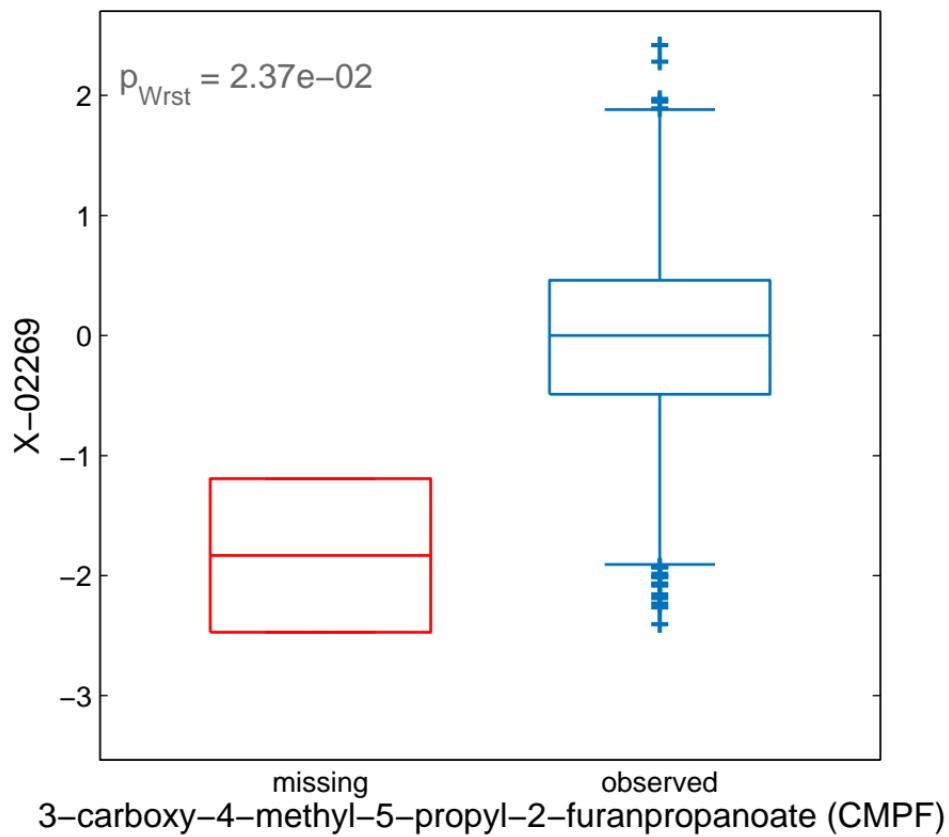


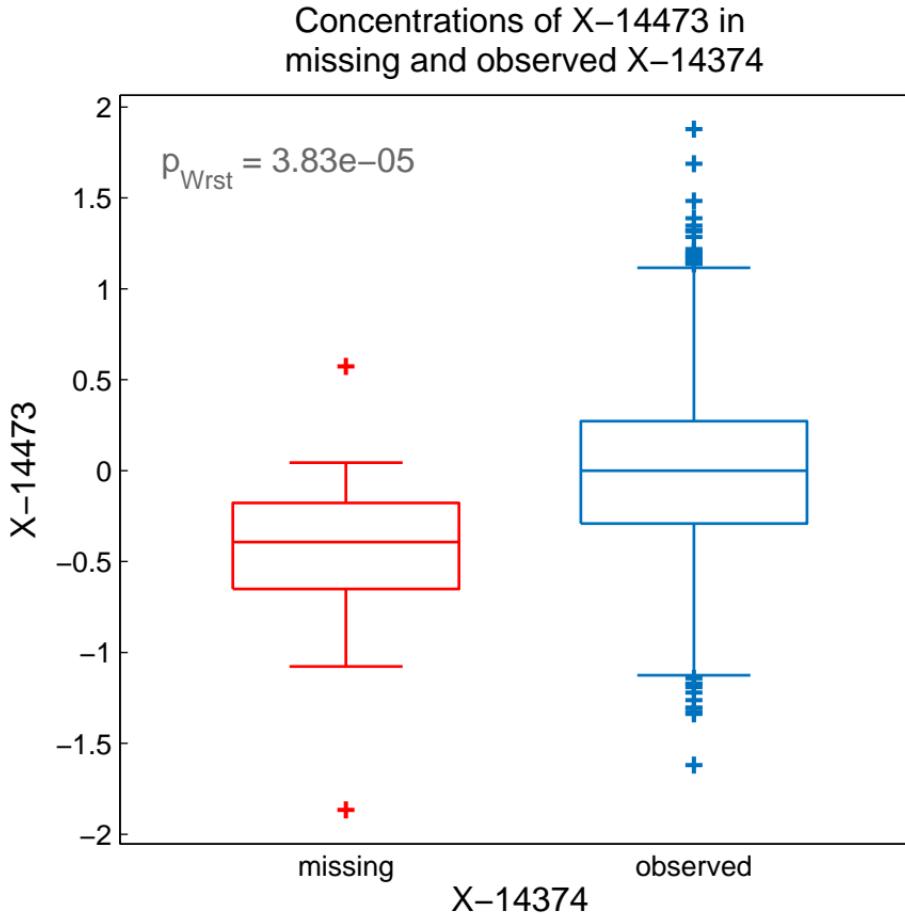
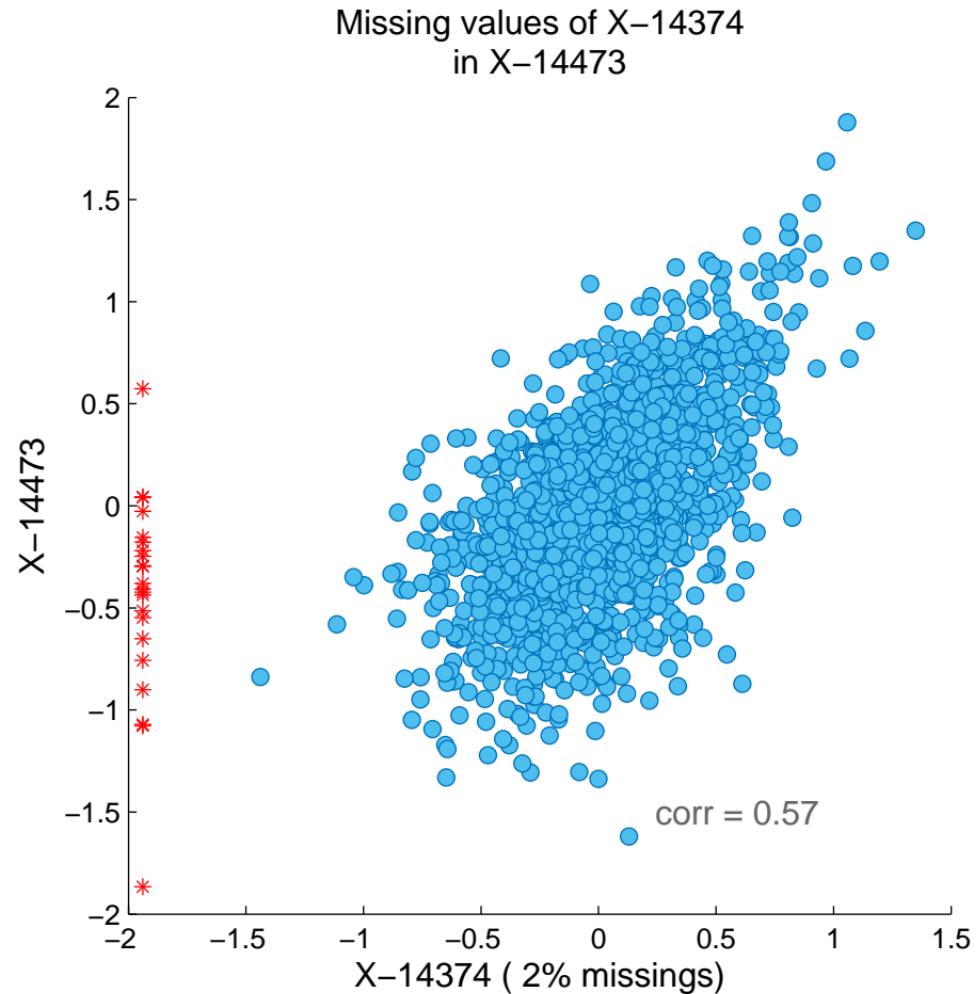


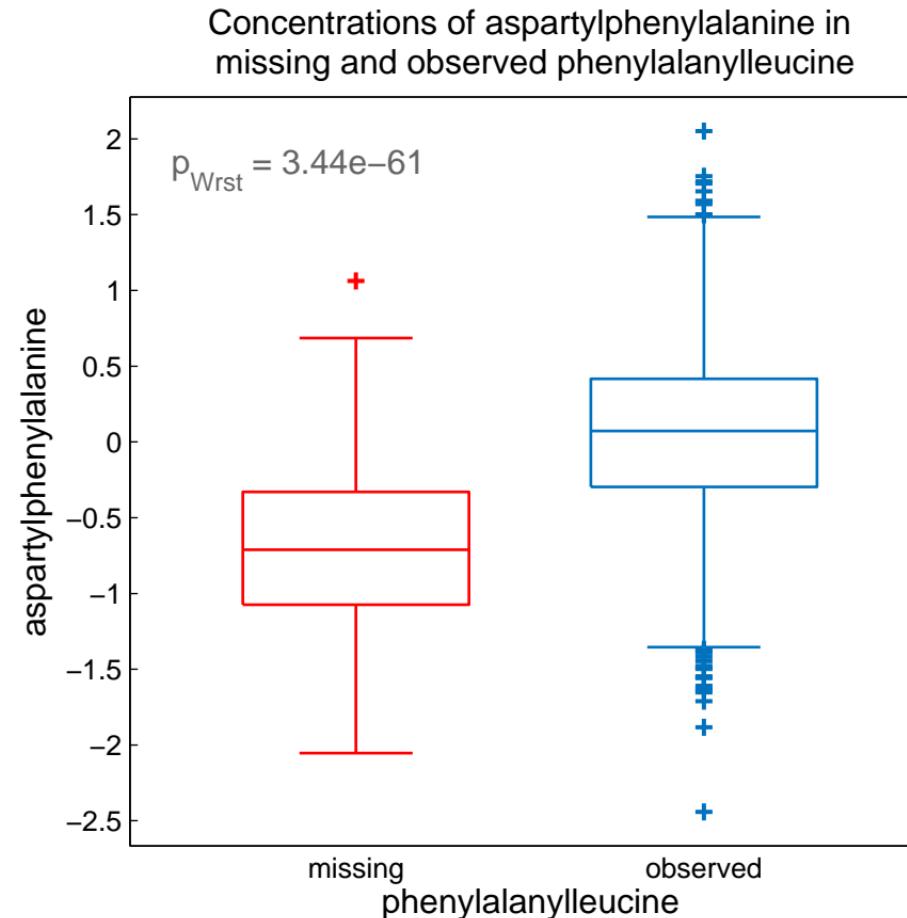
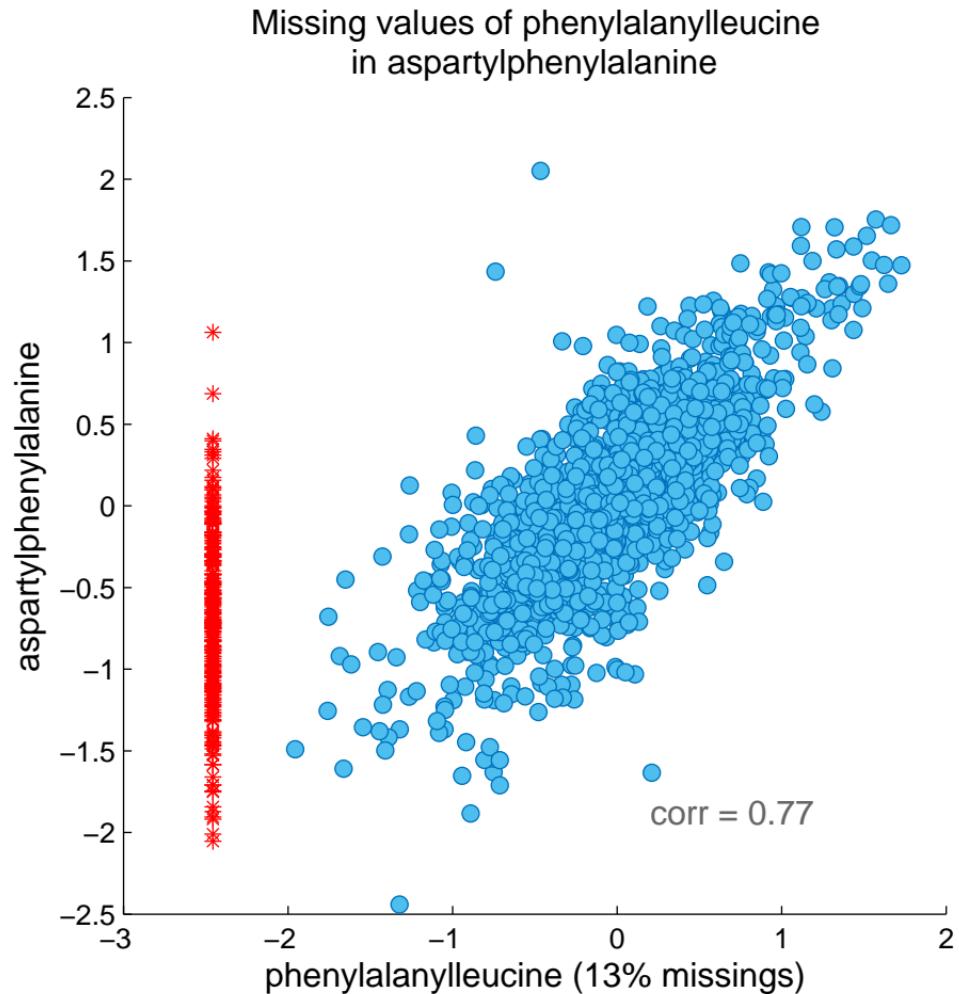
Missing values of 3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF)
in X-02269

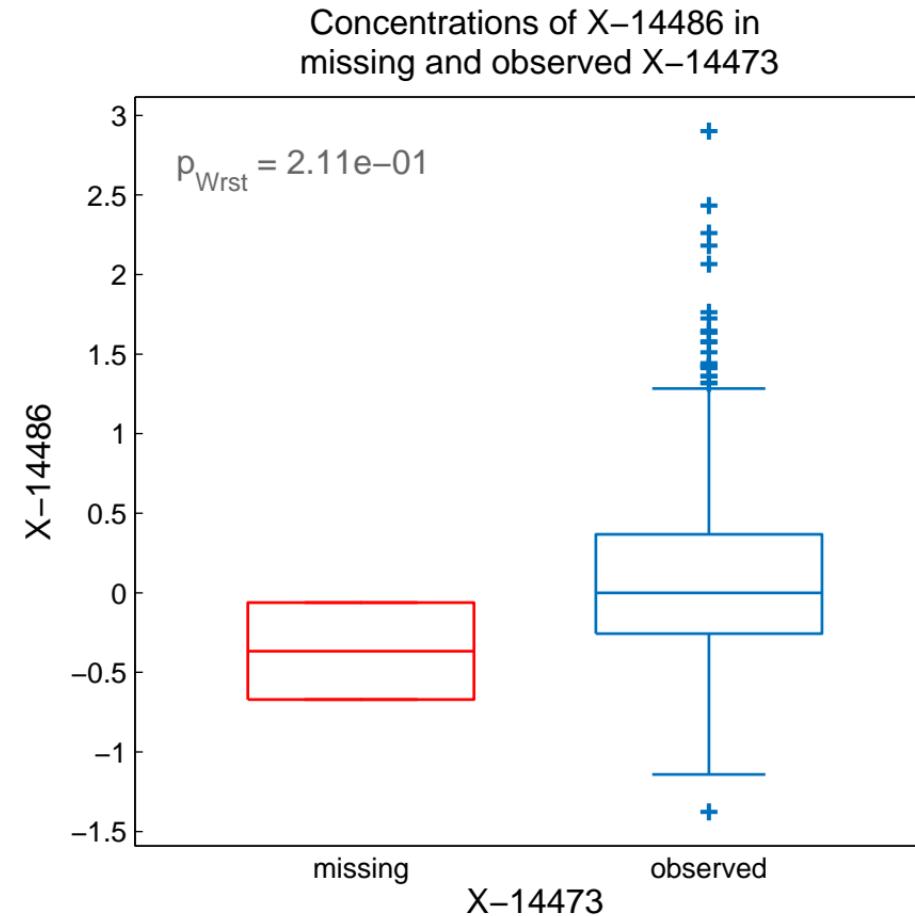
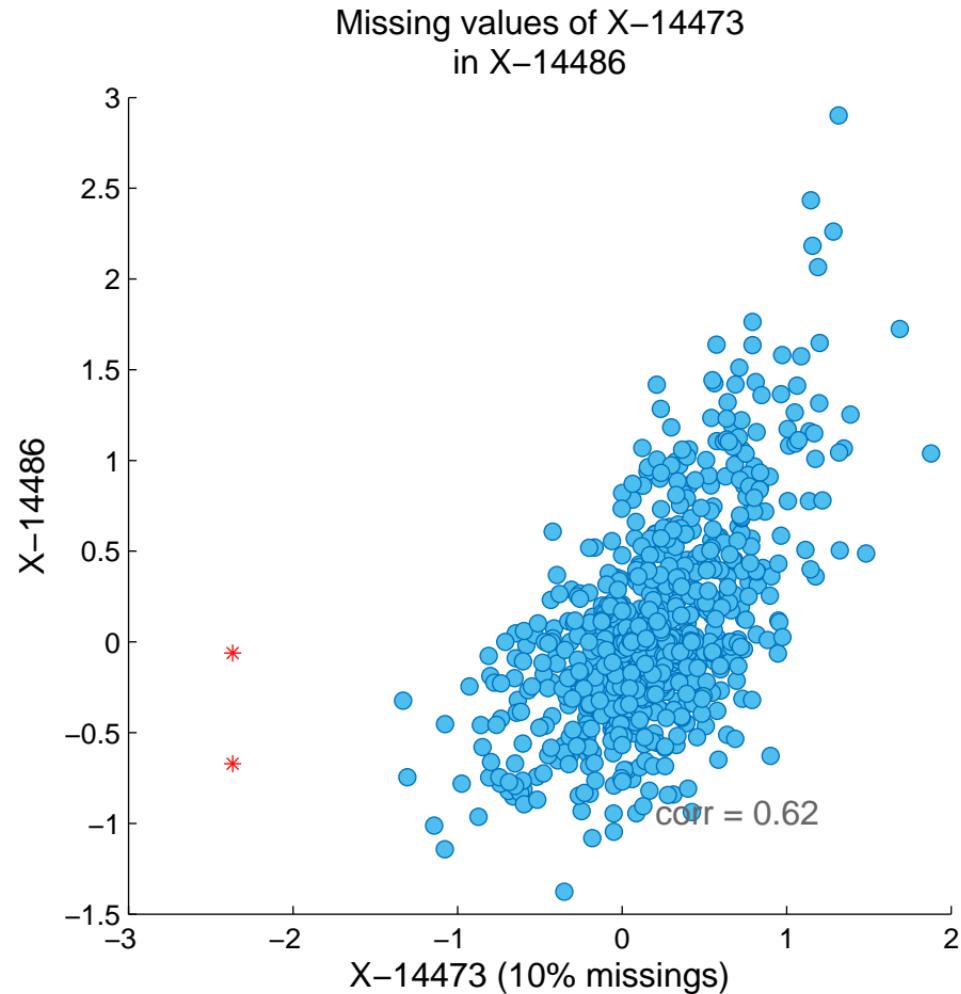


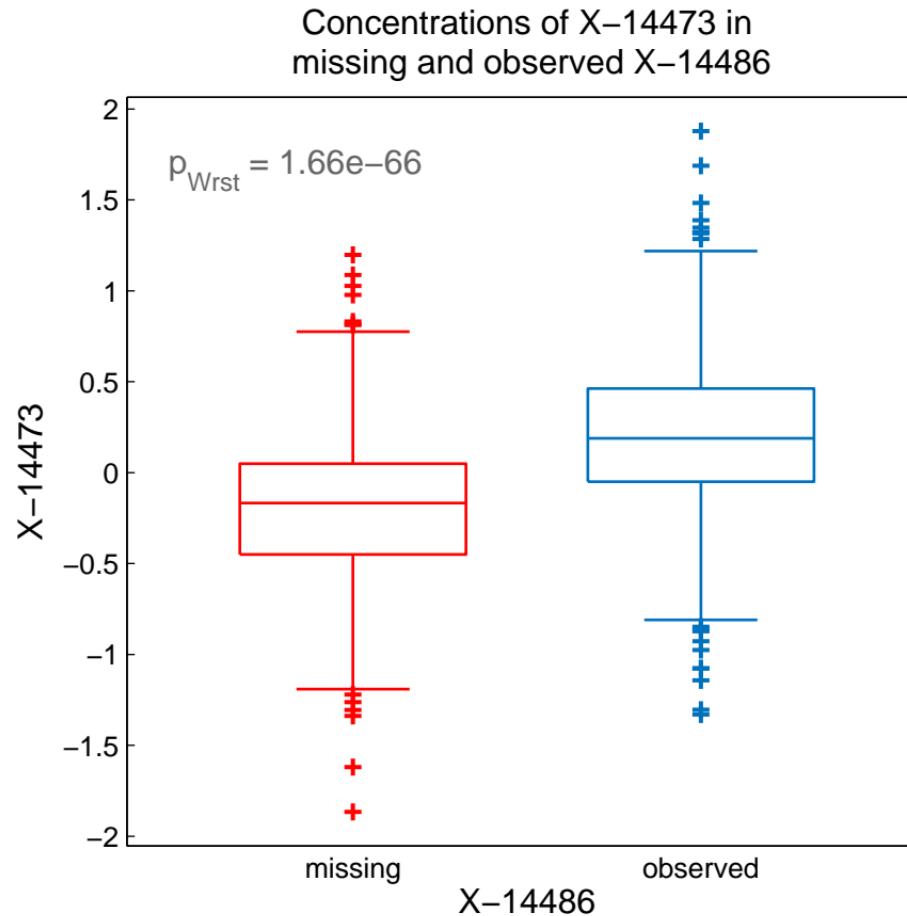
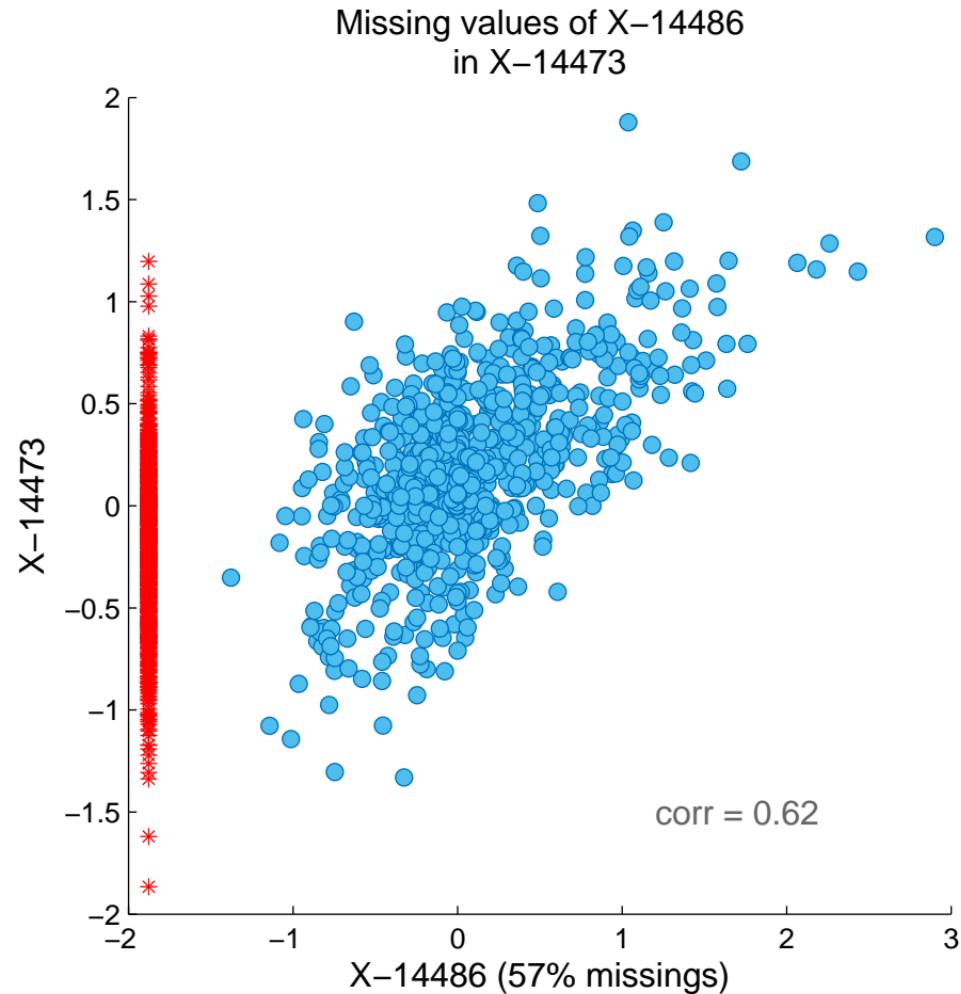
Concentrations of X-02269 in
missing and observed 3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF)

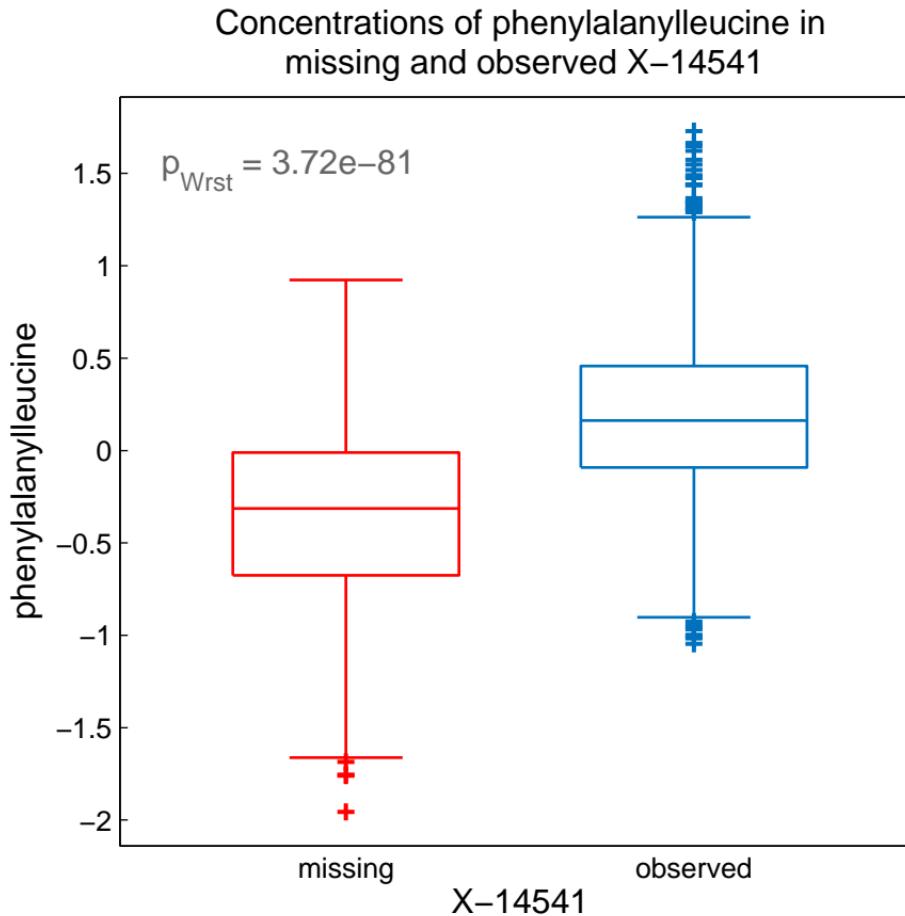
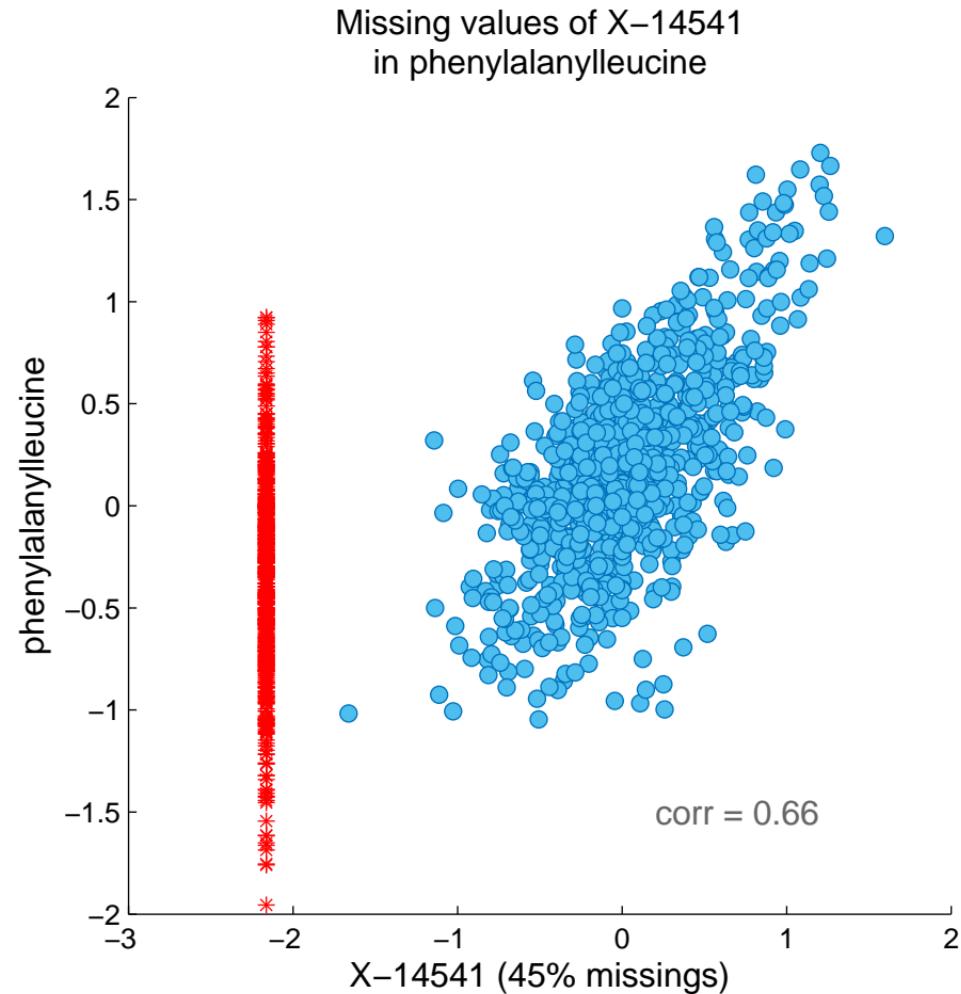


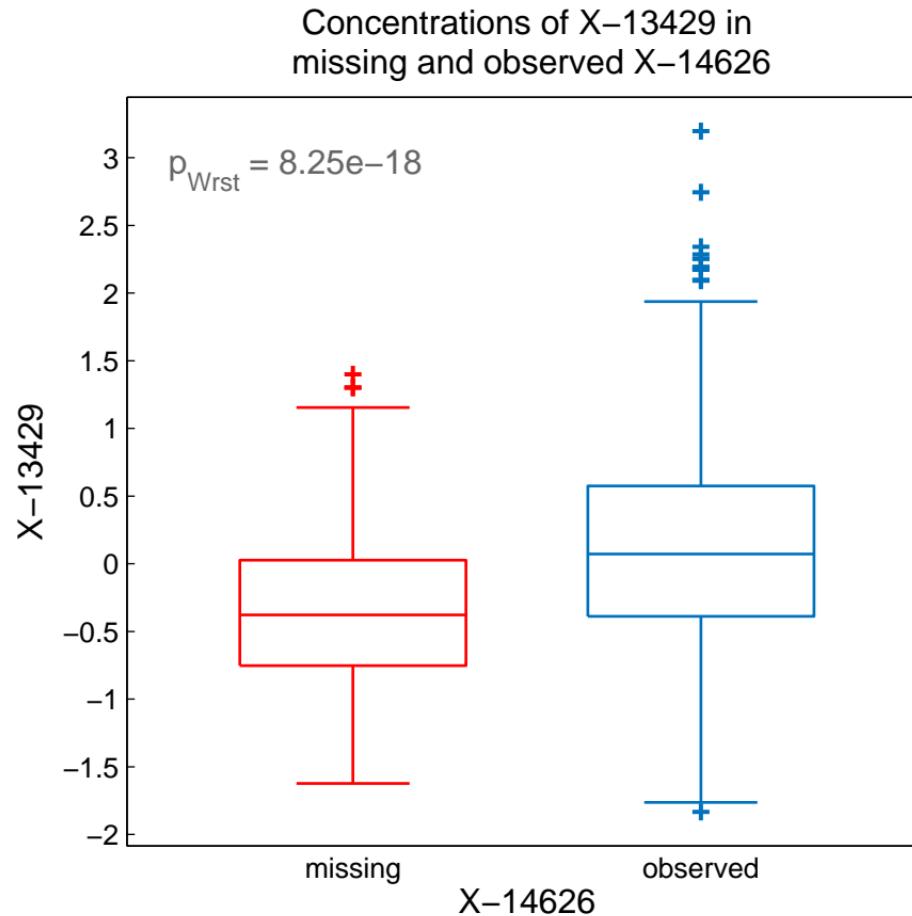
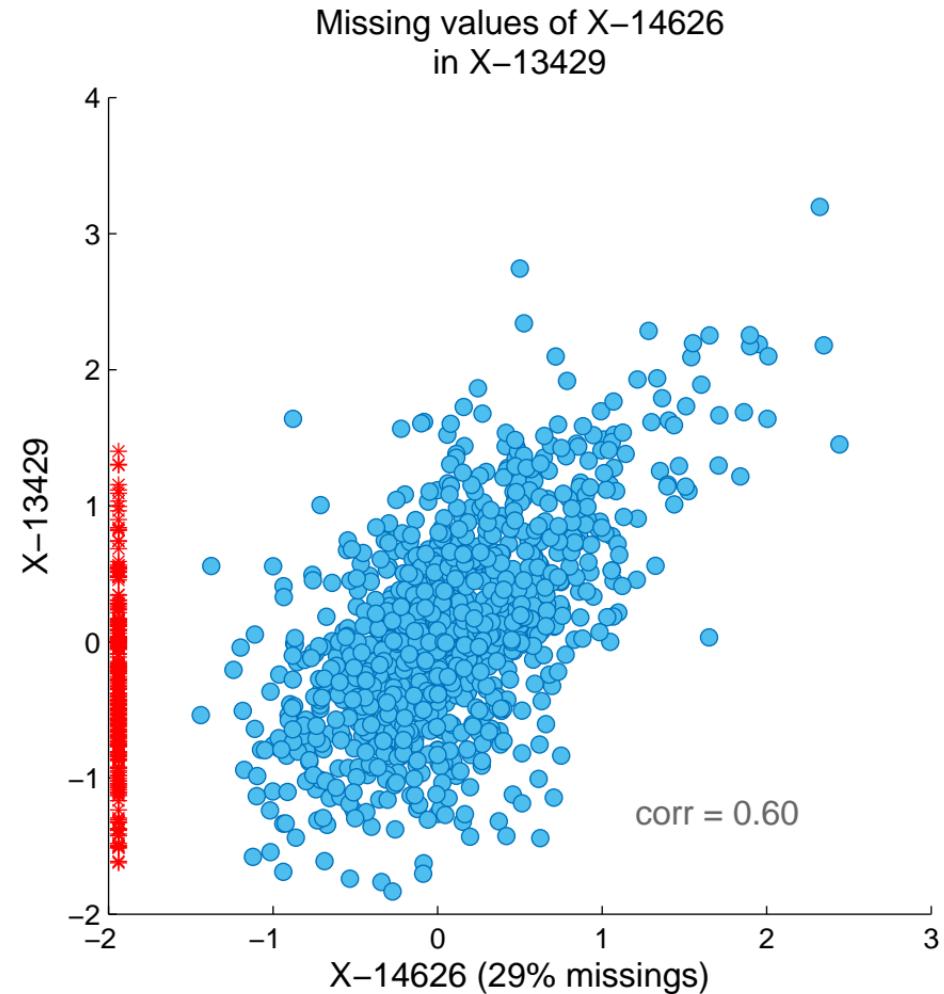


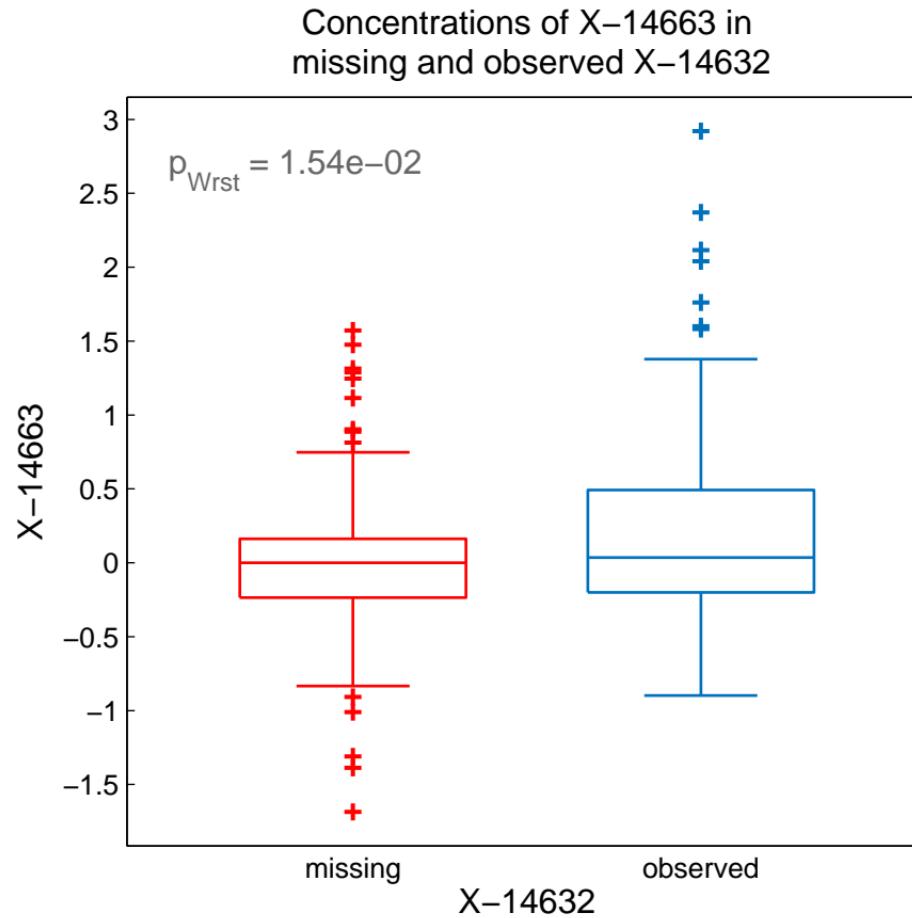
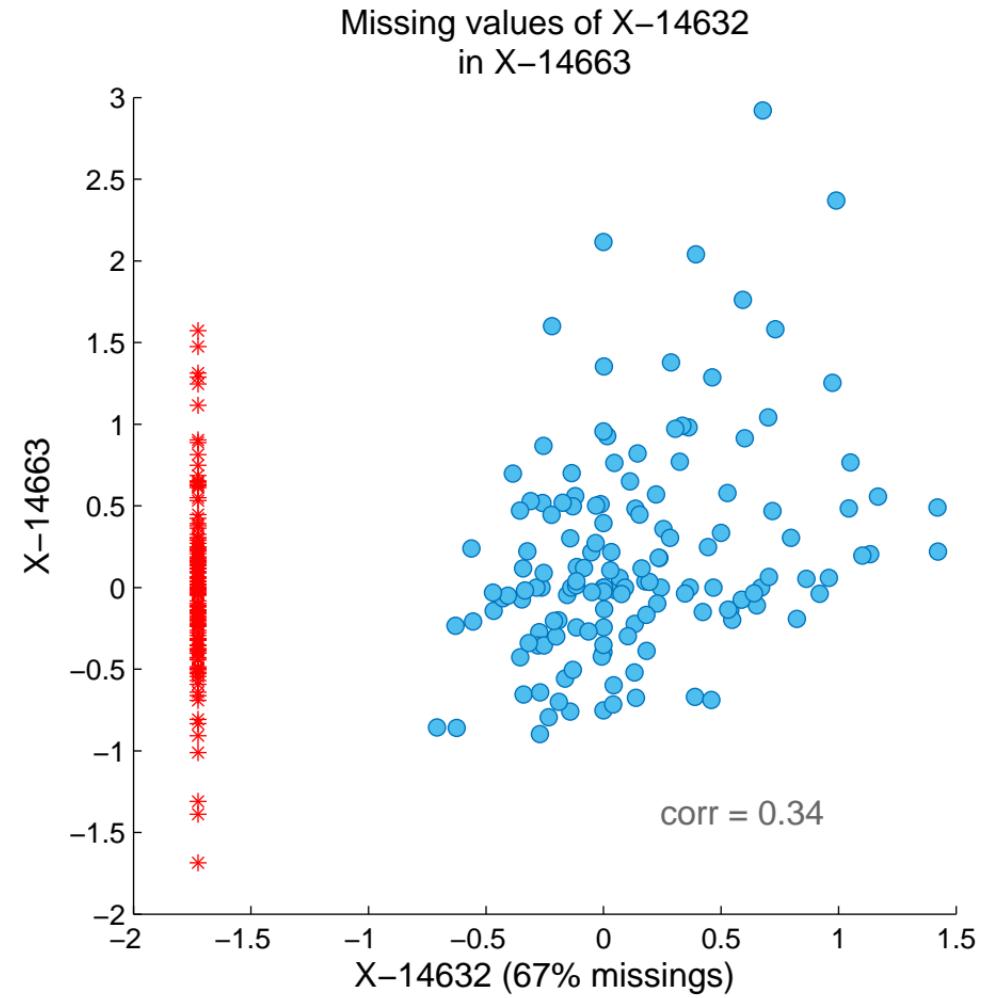


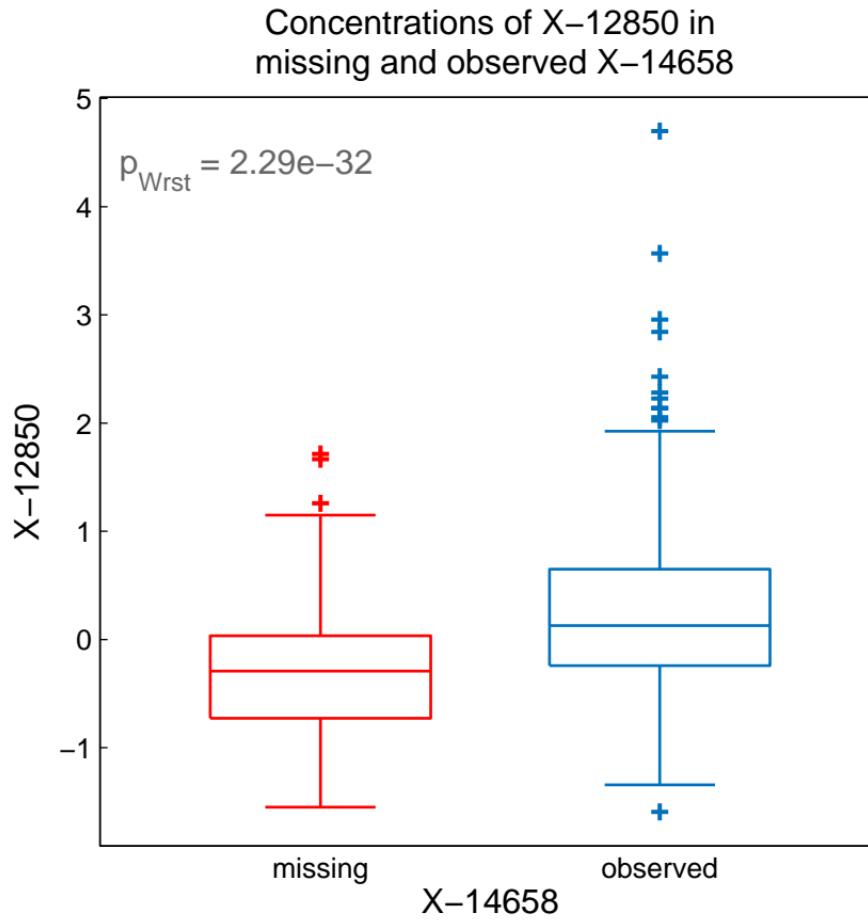
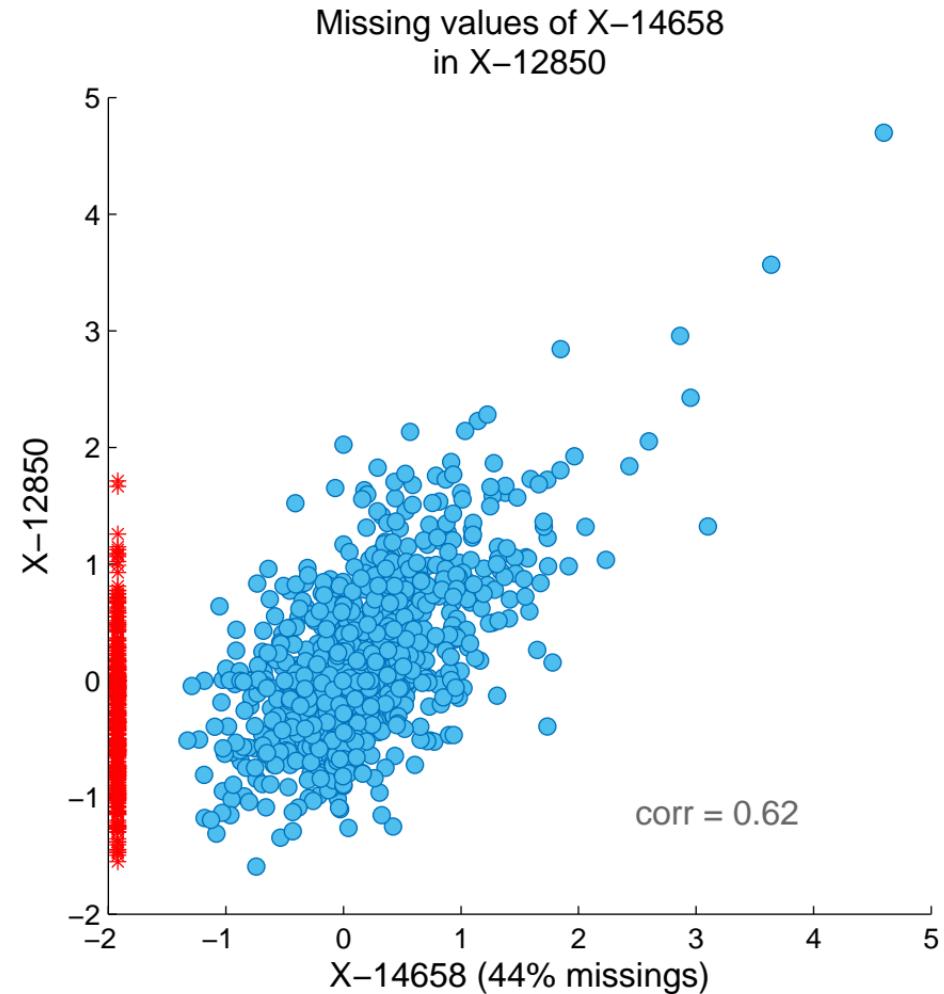


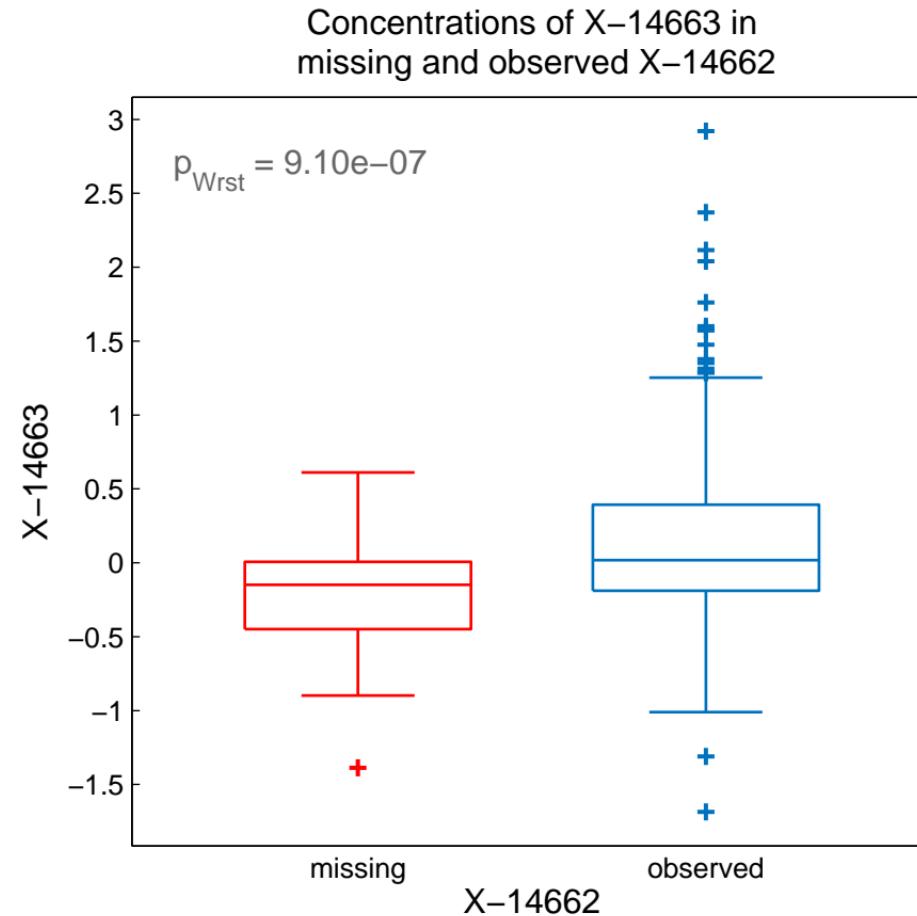
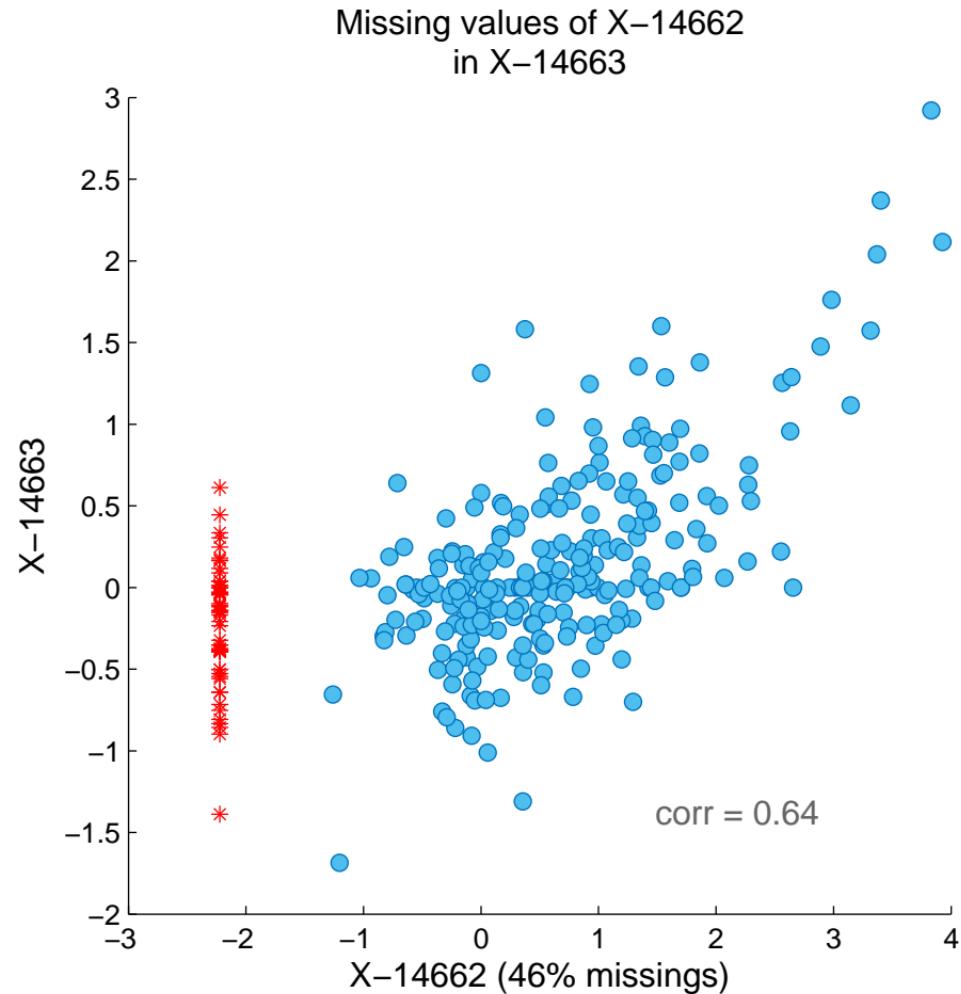


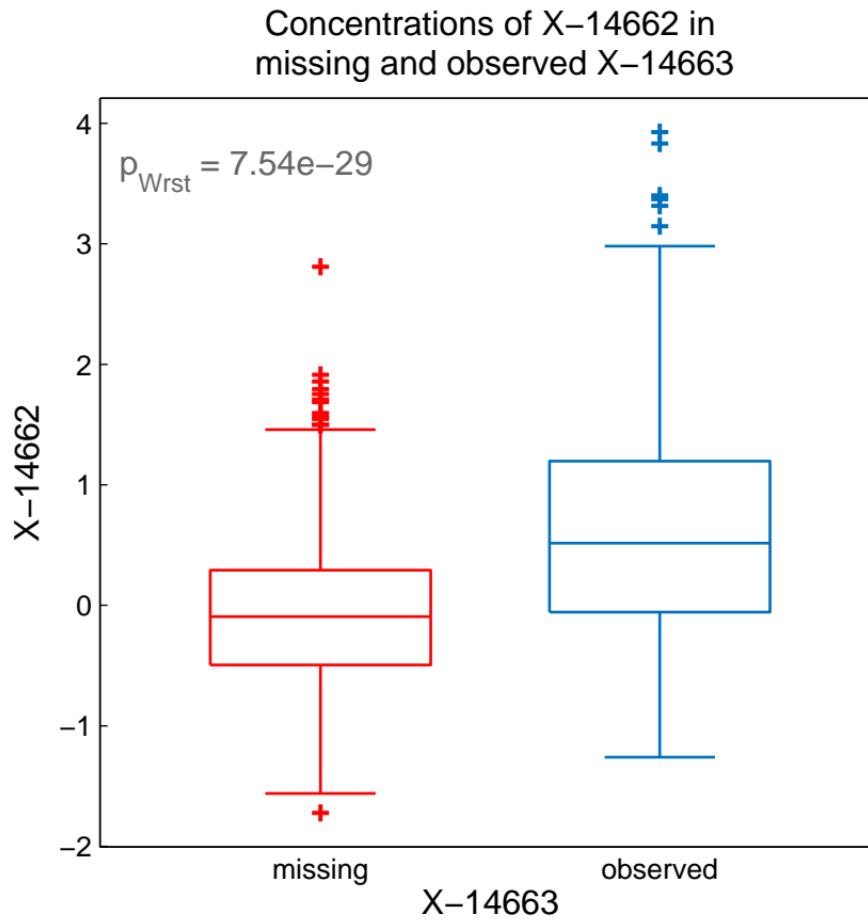
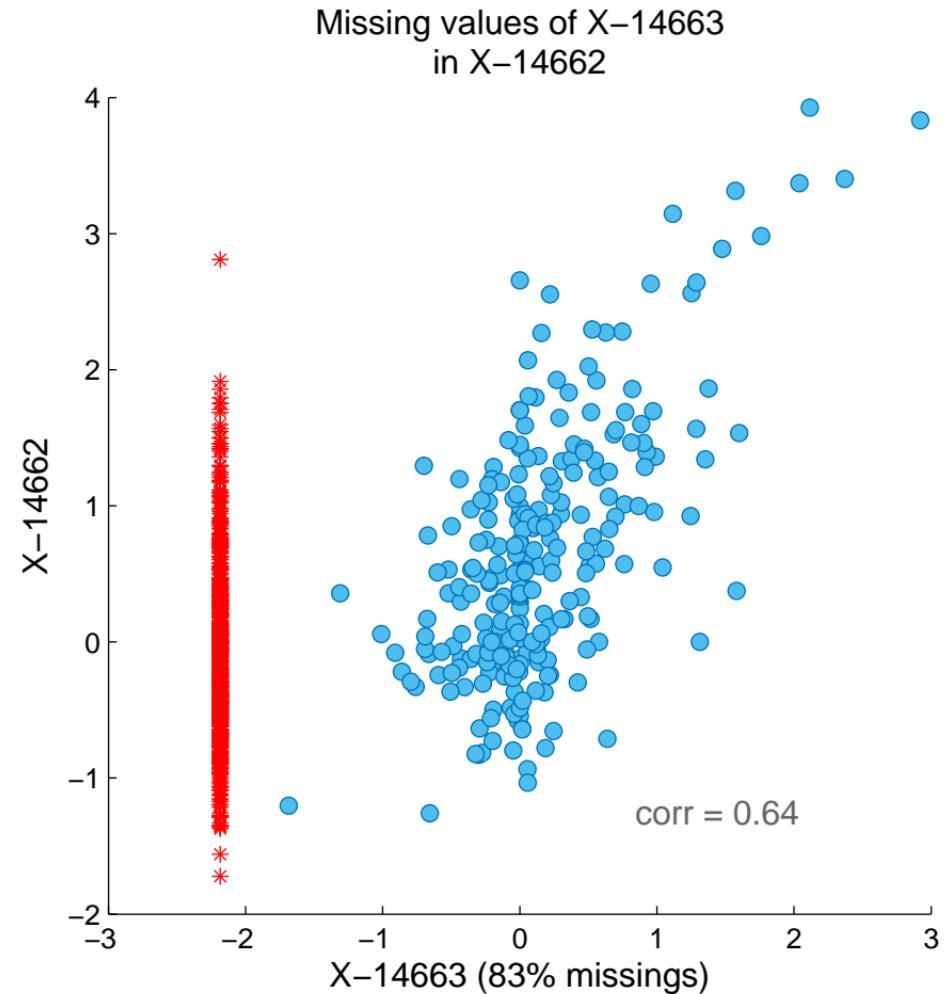


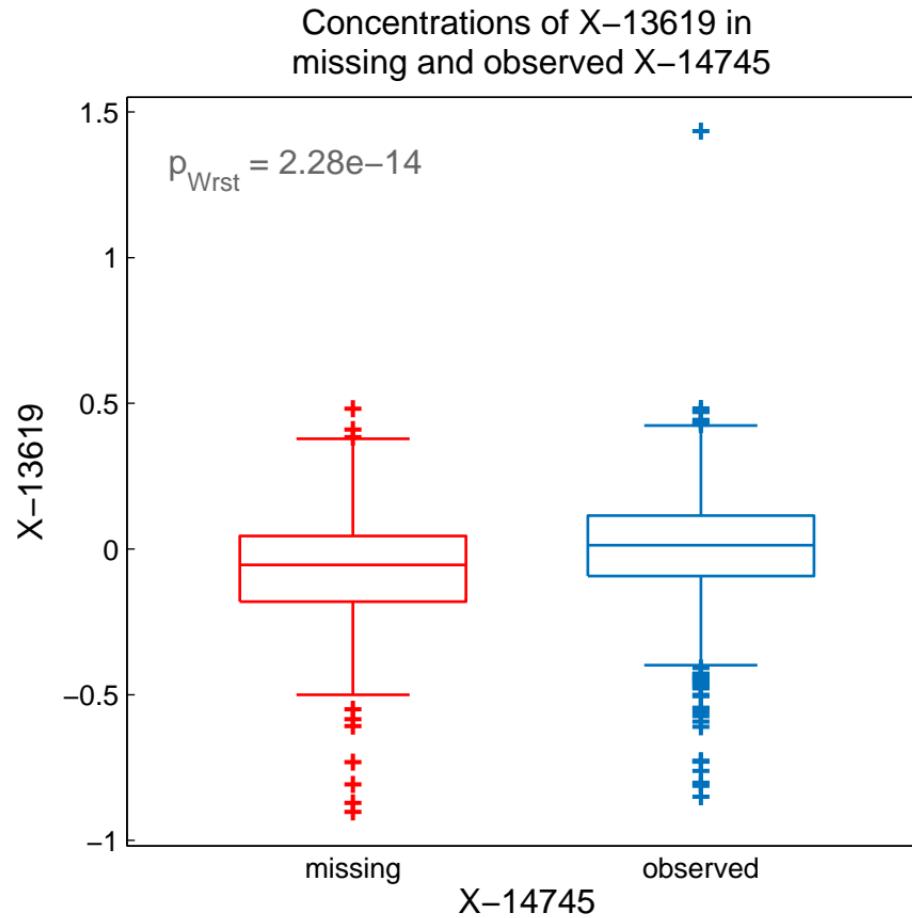
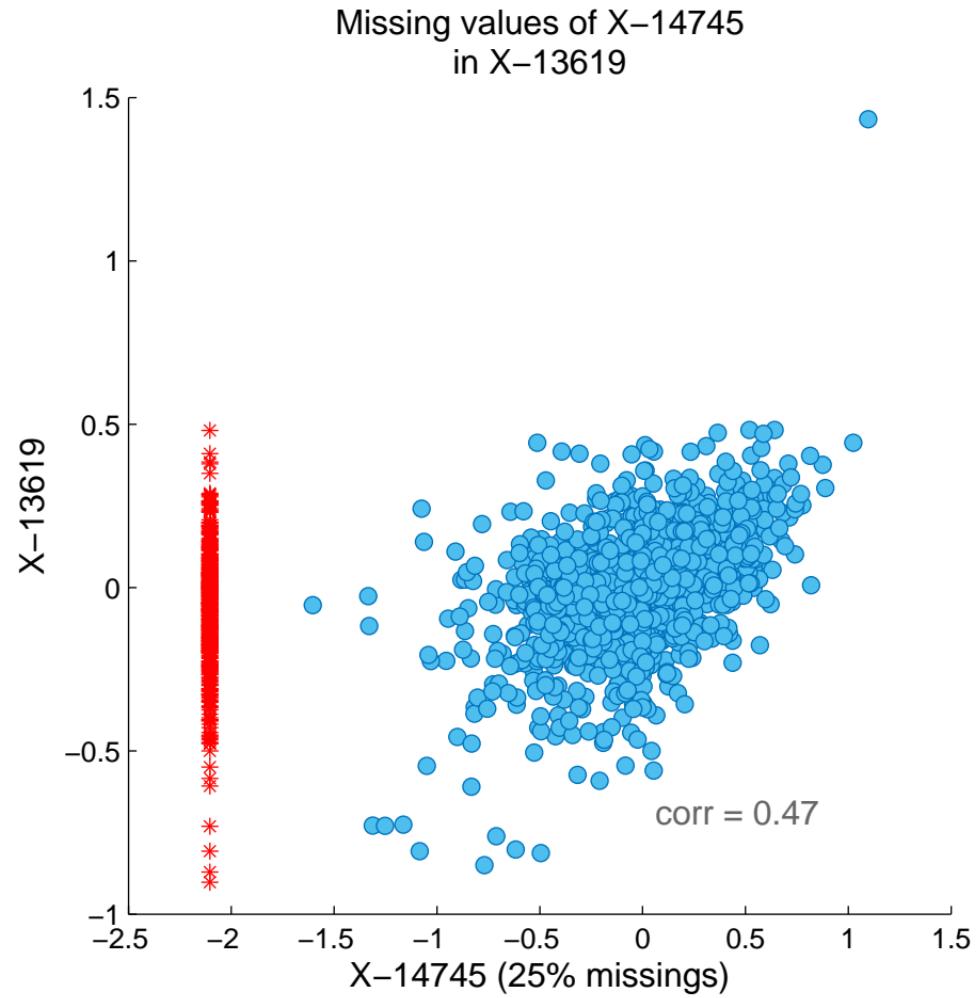




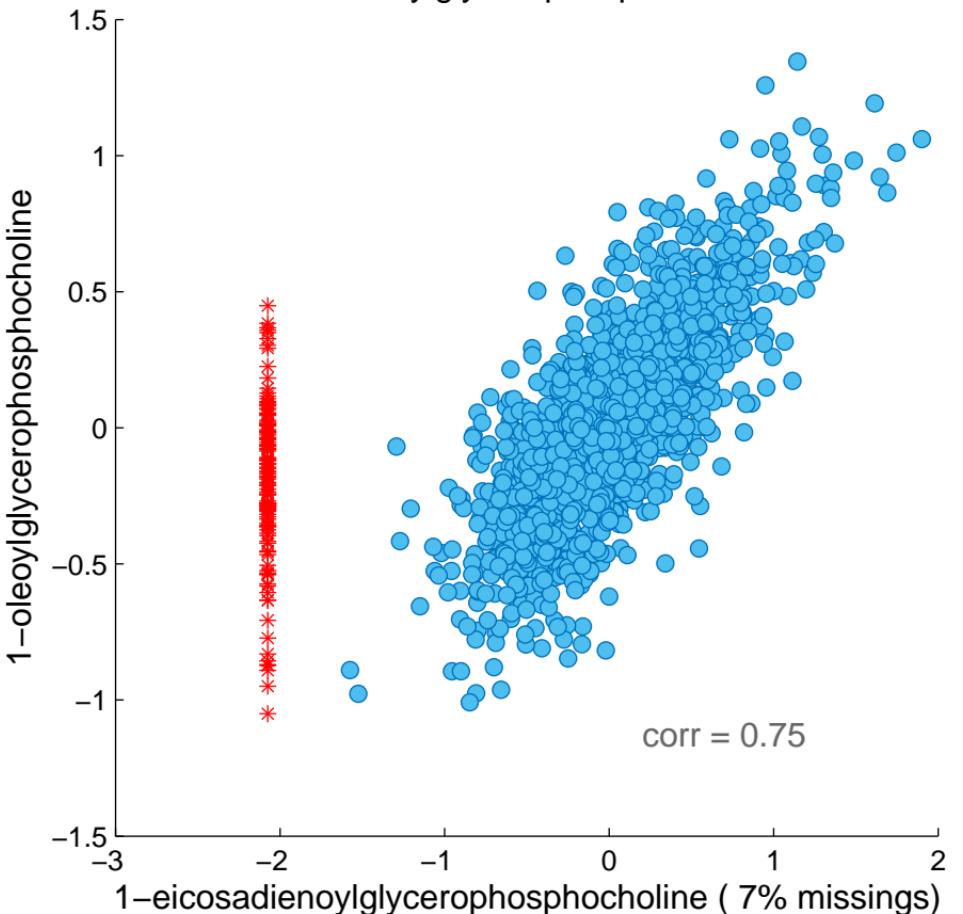




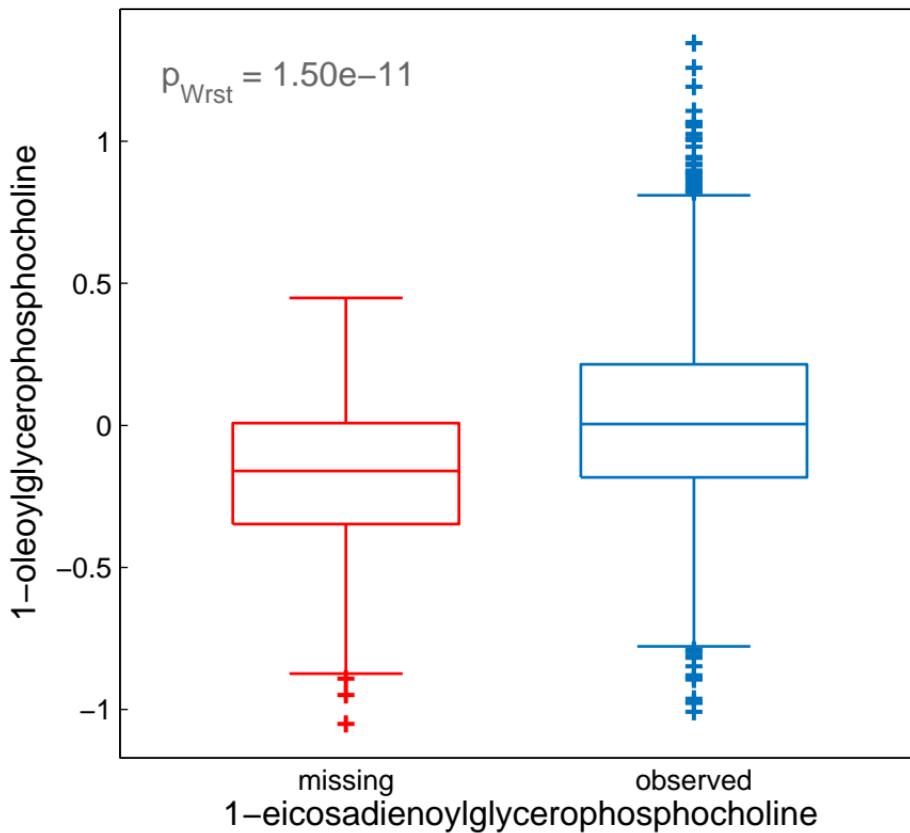


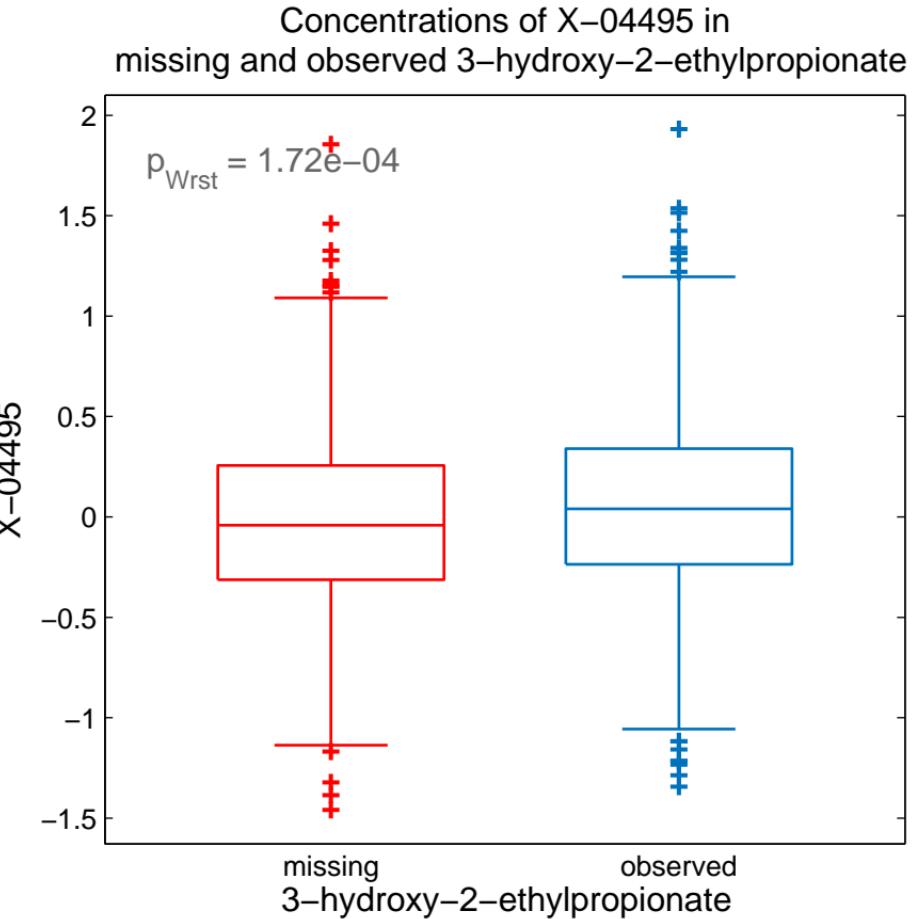
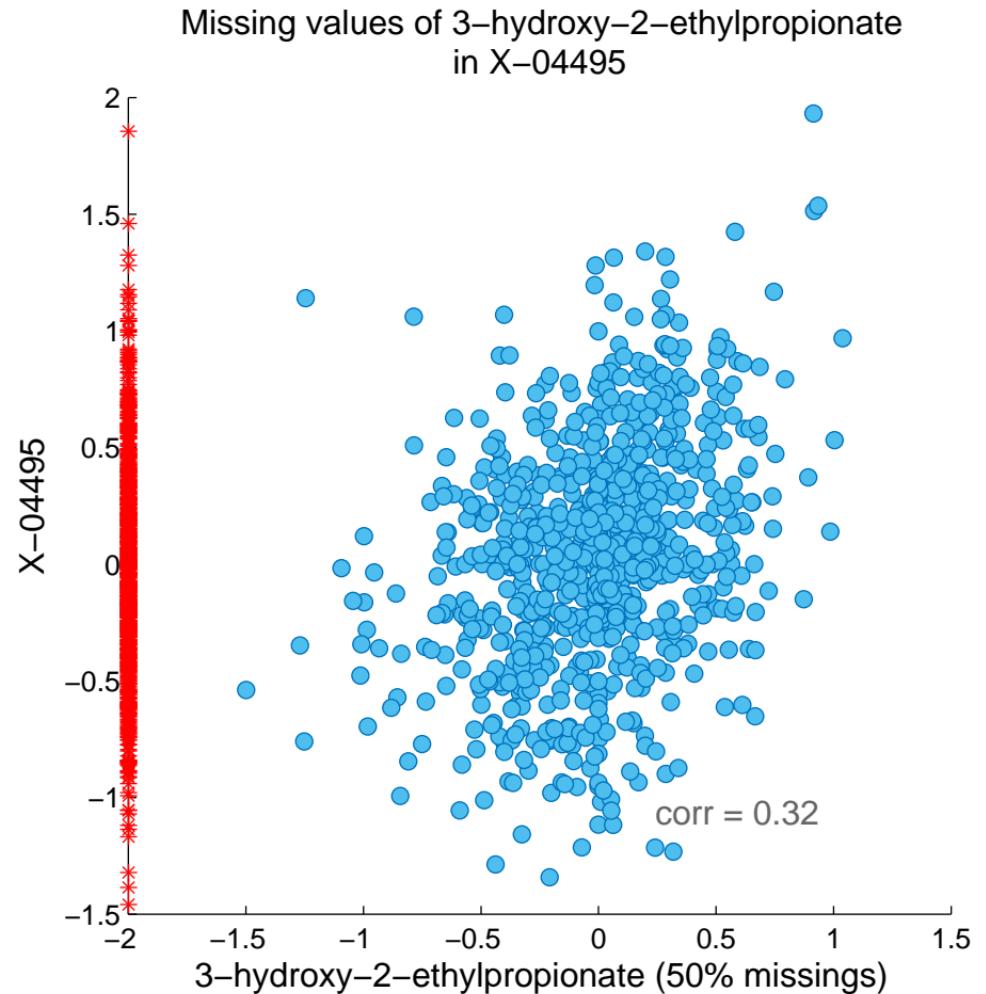


Missing values of 1-eicosadienoylglycerophosphocholine
in 1-oleoylglycerophosphocholine

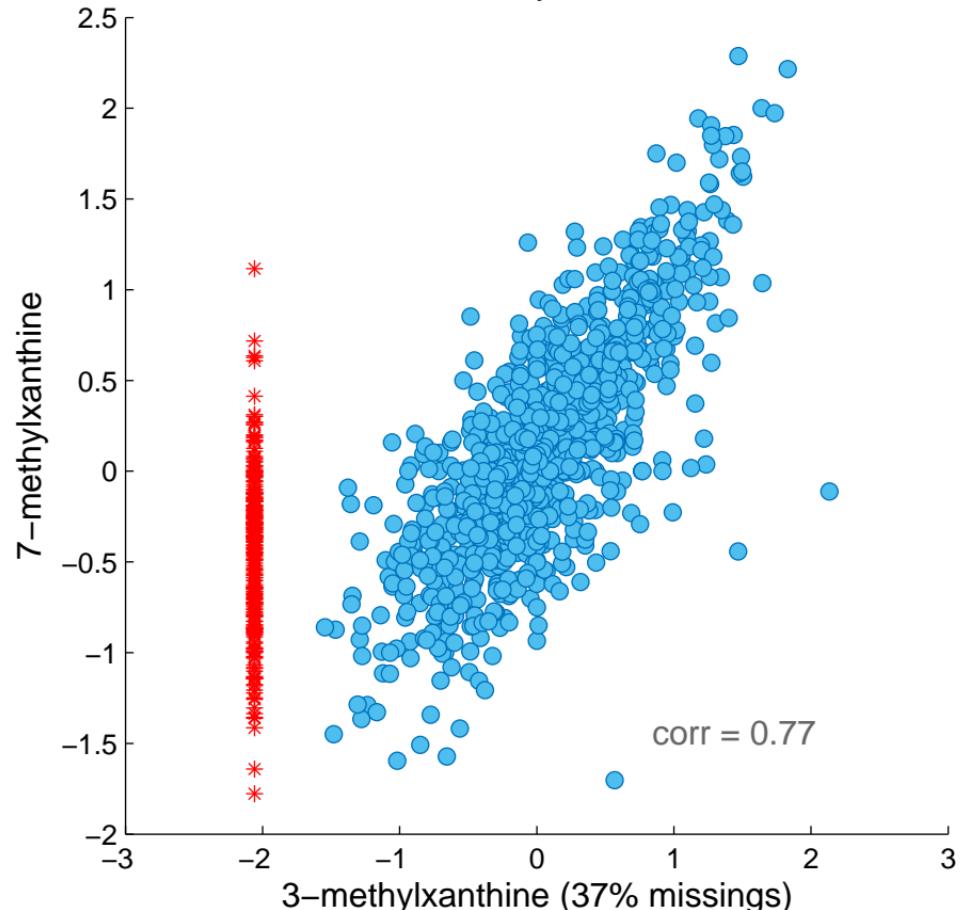


Concentrations of 1-oleoylglycerophosphocholine in
missing and observed 1-eicosadienoylglycerophosphocholine

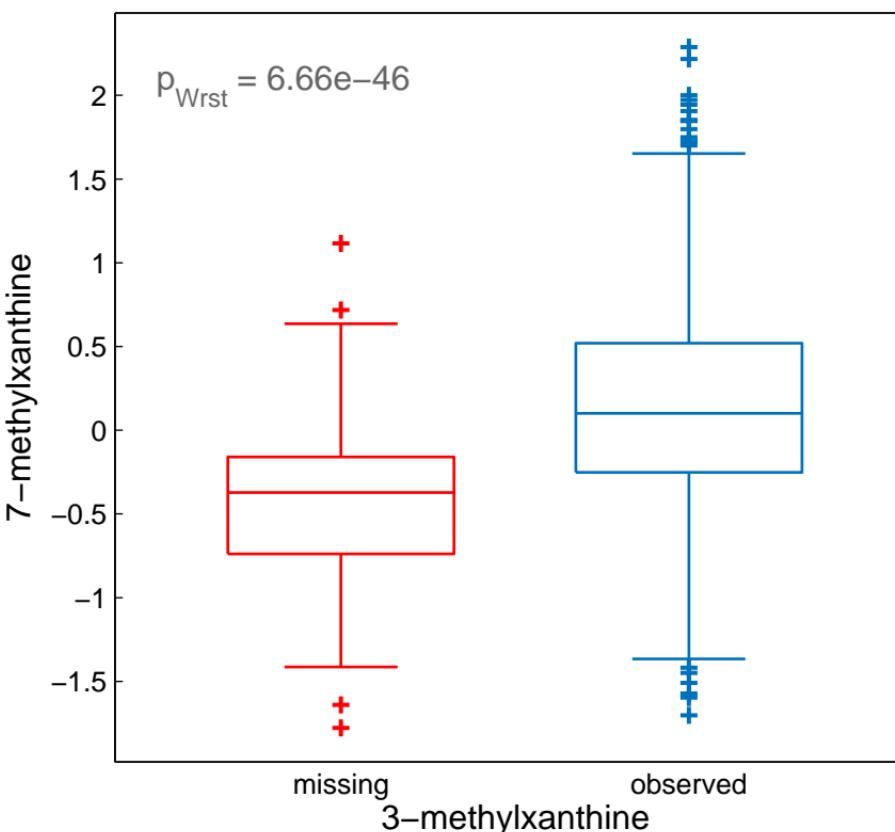


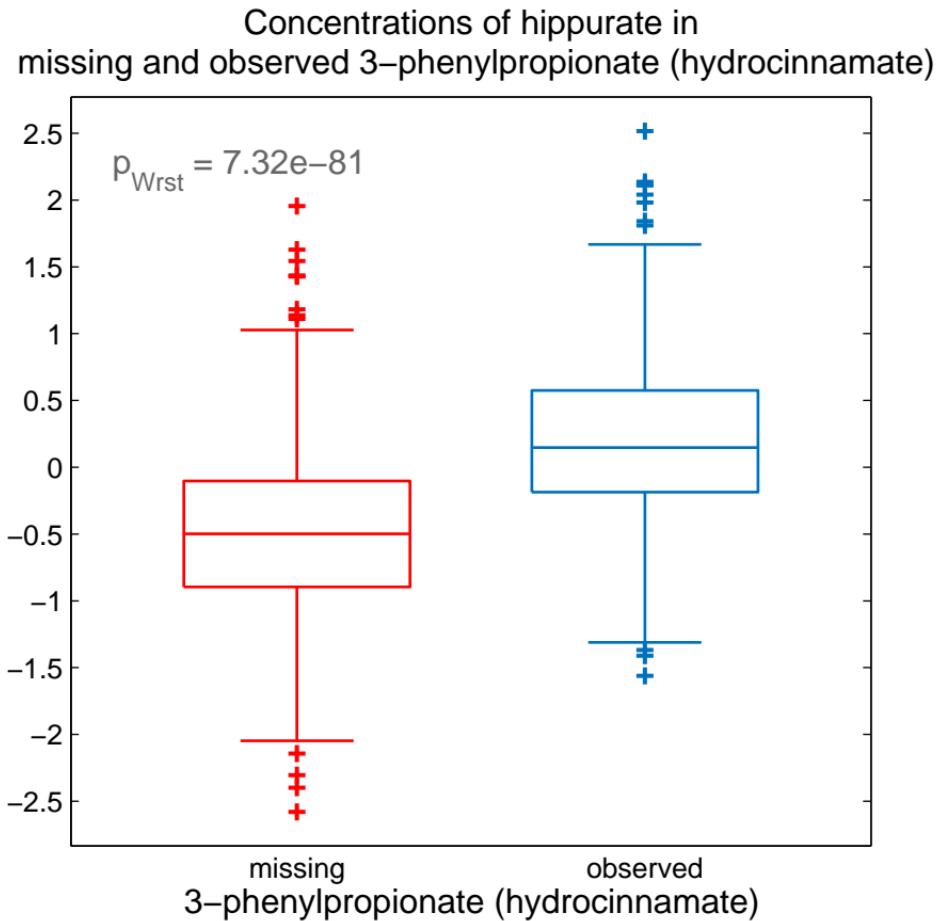
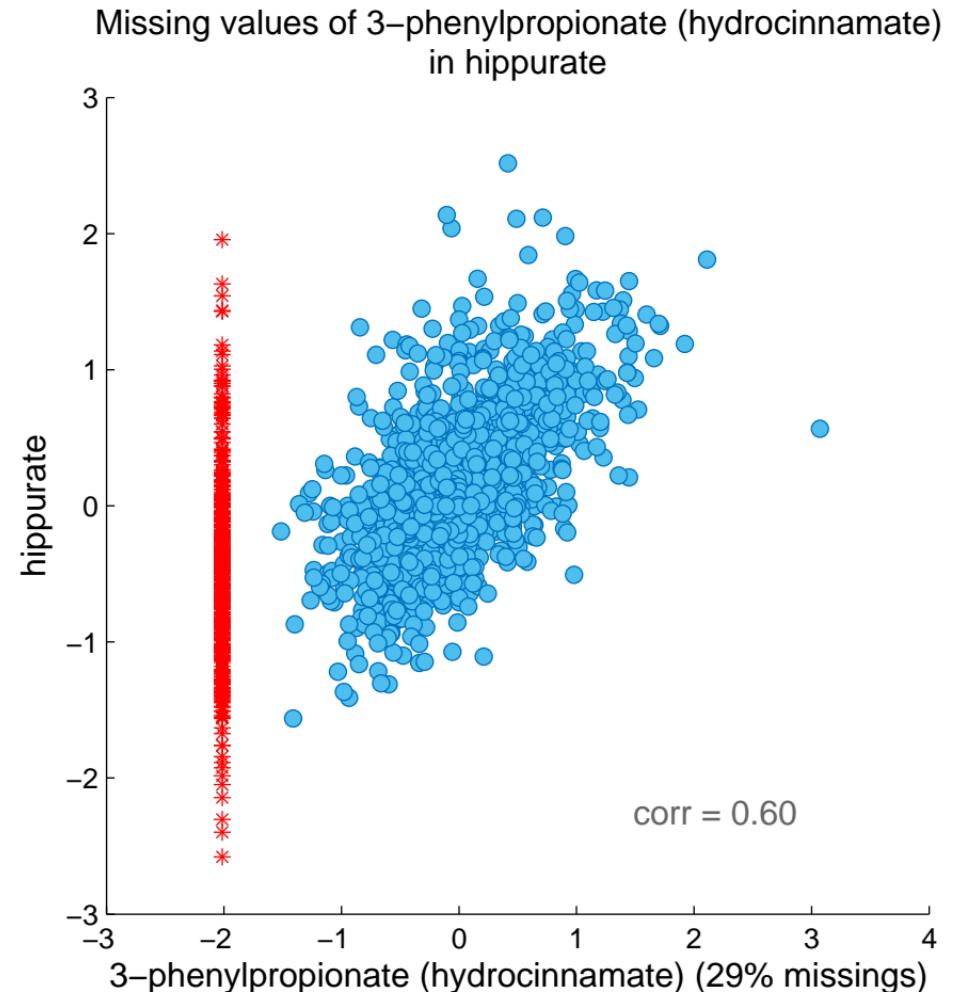


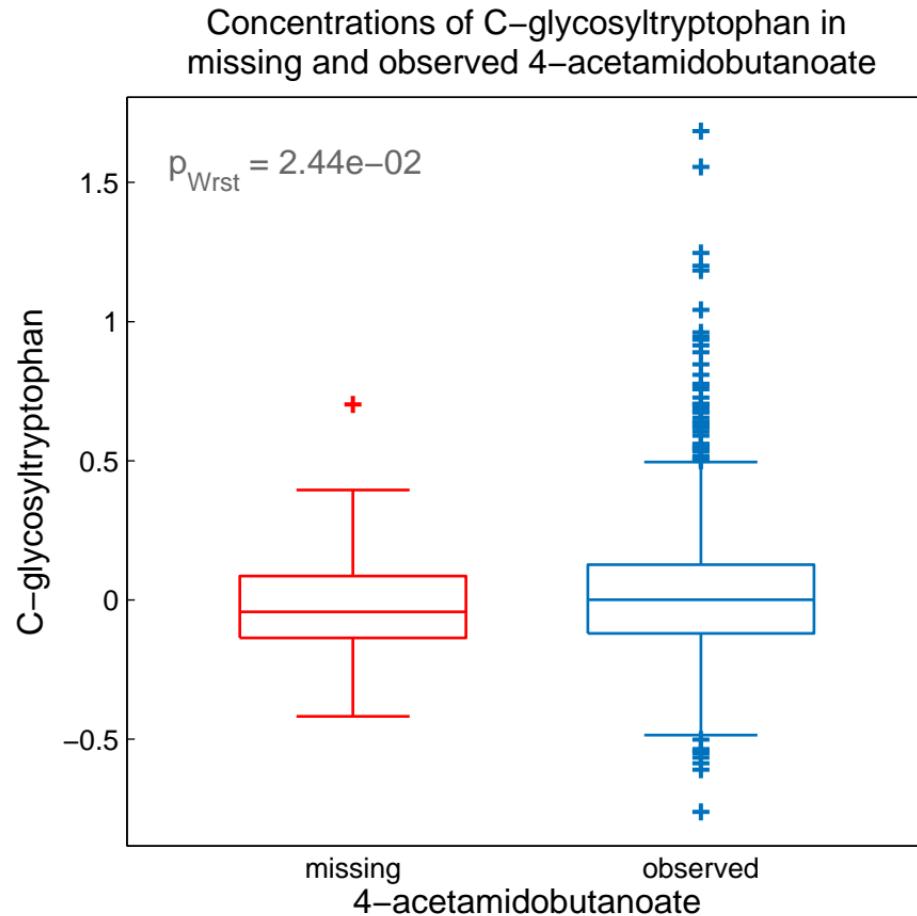
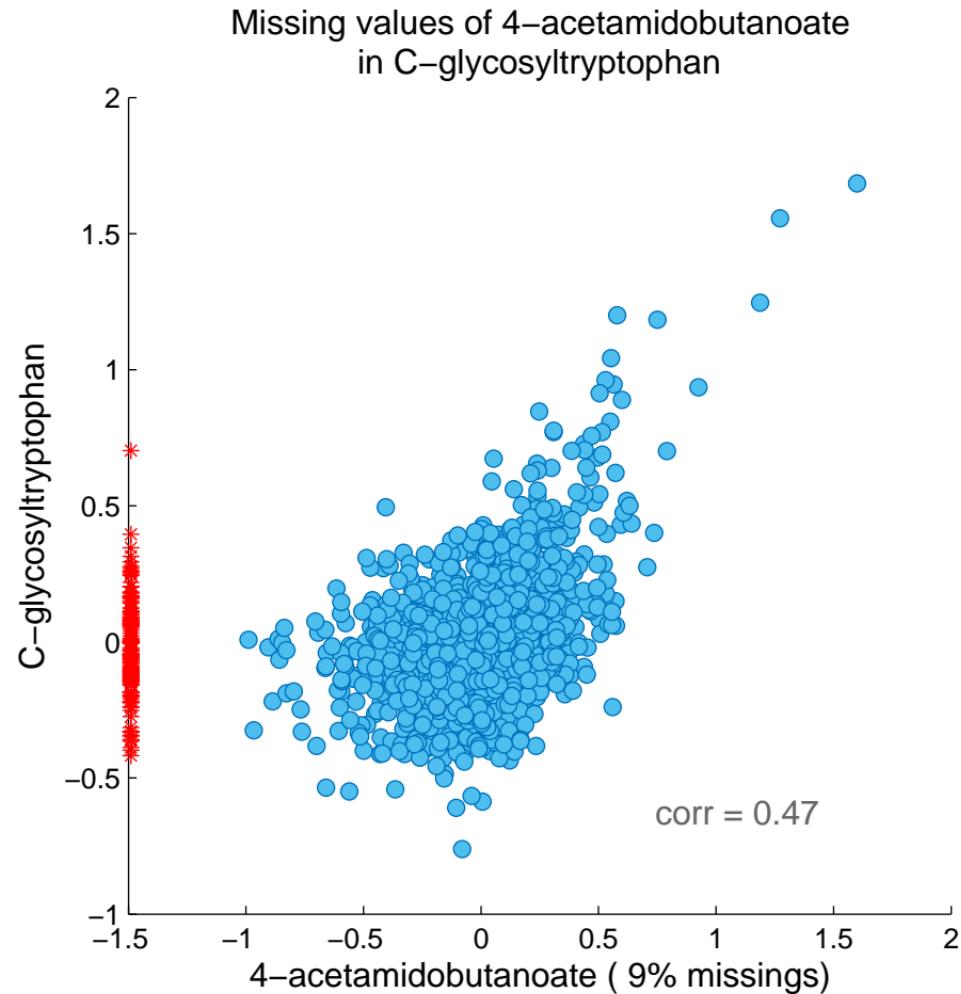
Missing values of 3-methylxanthine
in 7-methylxanthine

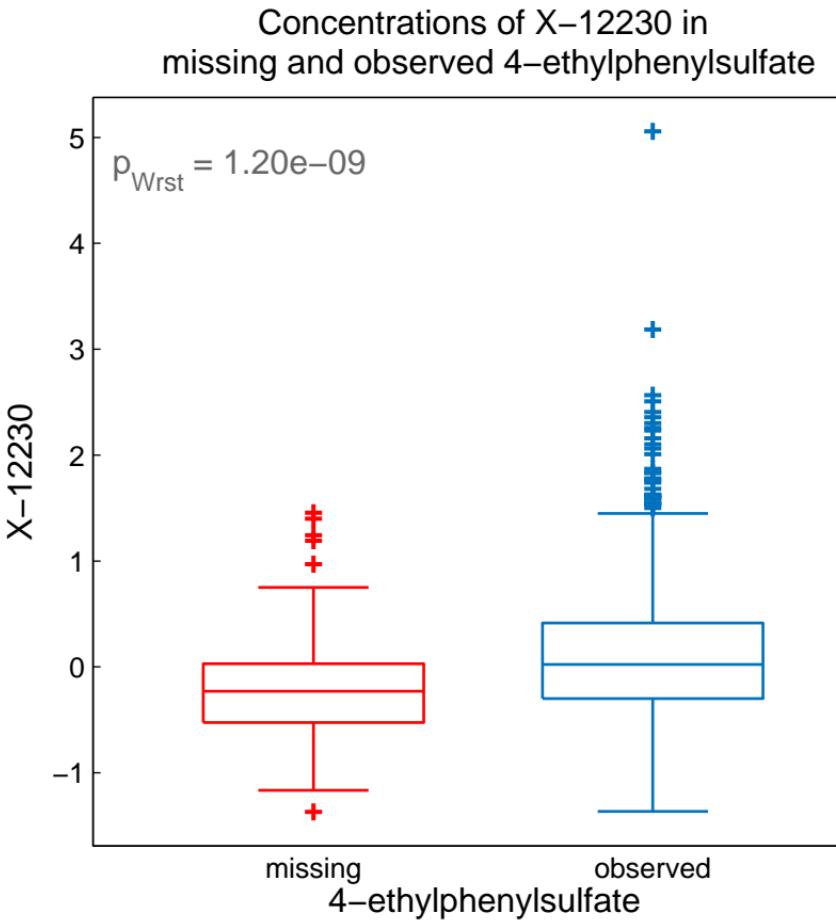
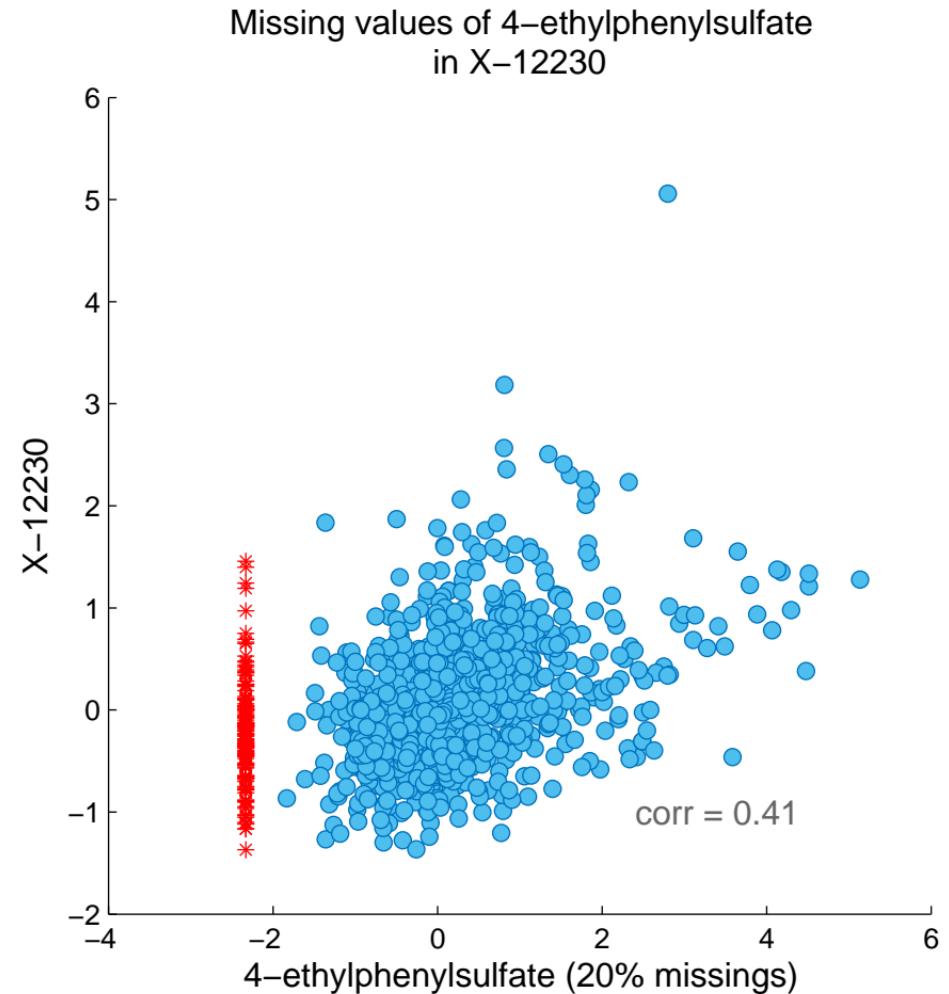


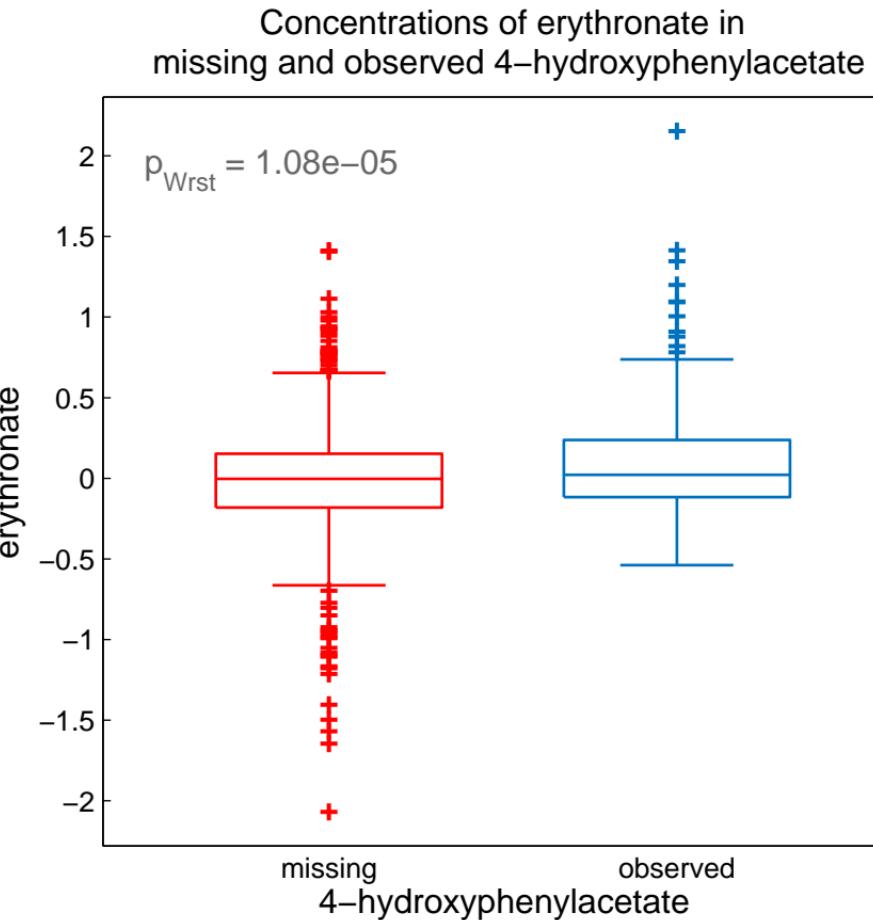
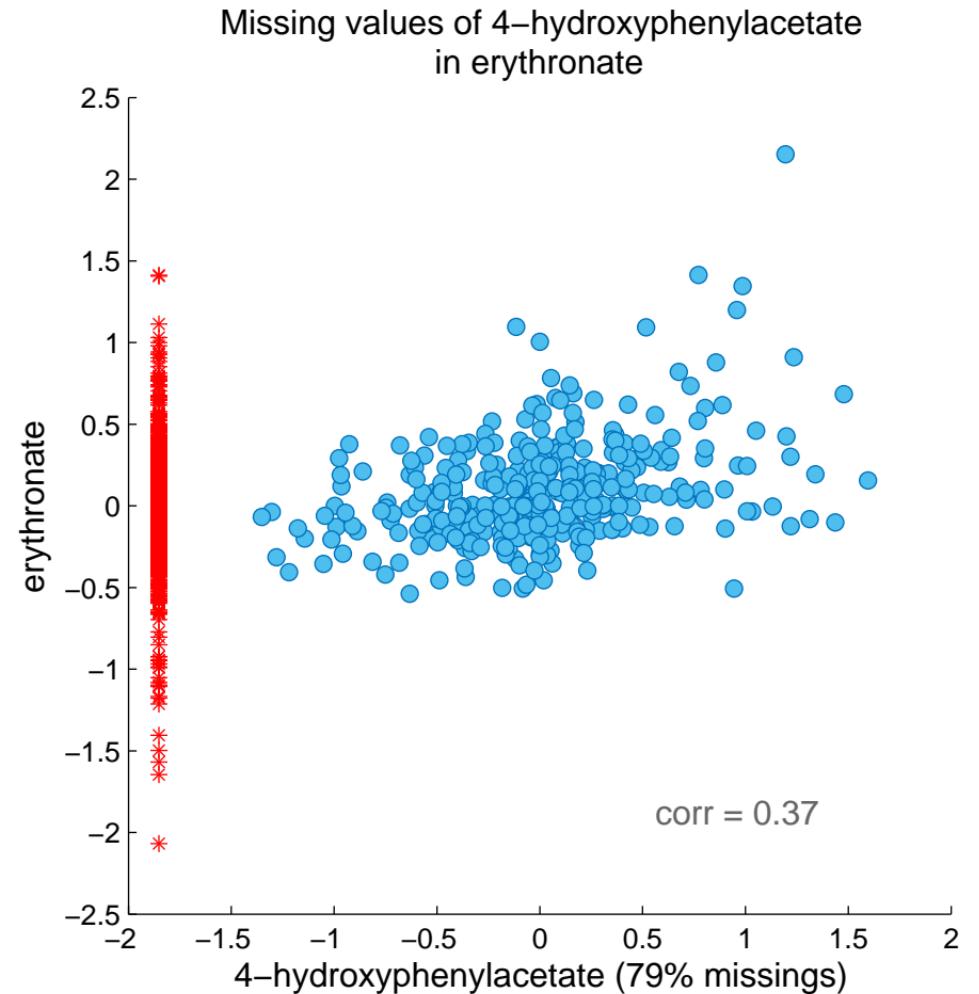
Concentrations of 7-methylxanthine in
missing and observed 3-methylxanthine

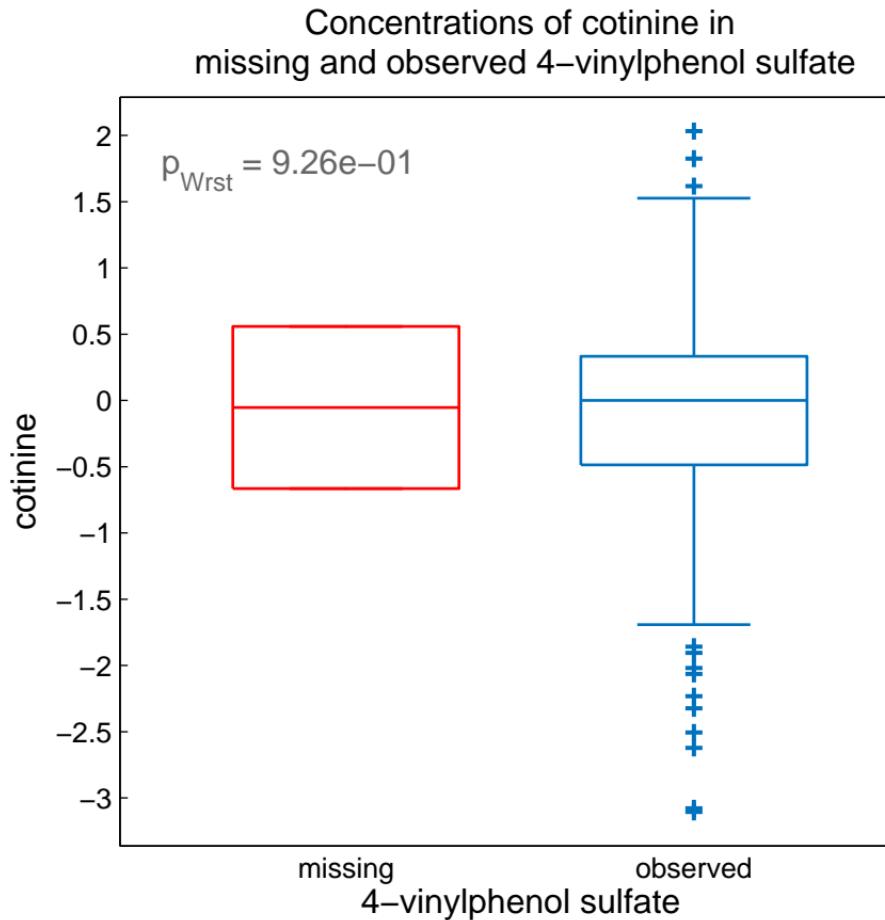
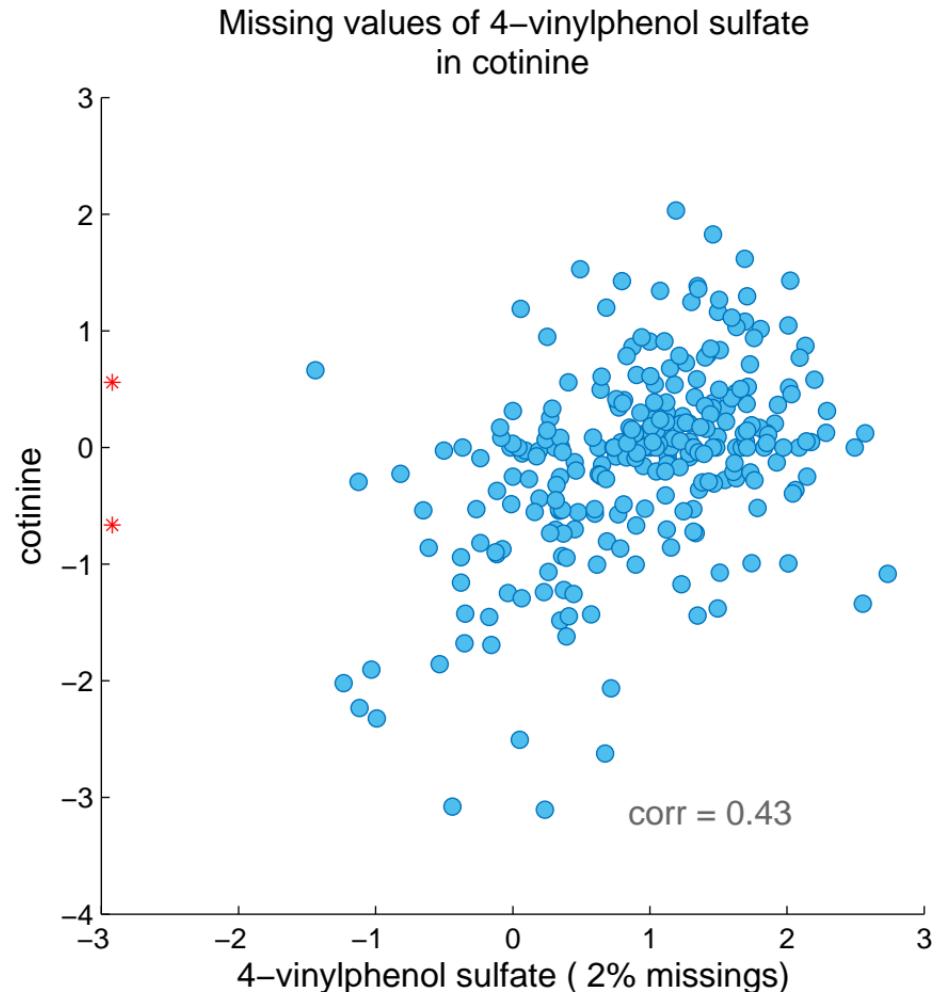


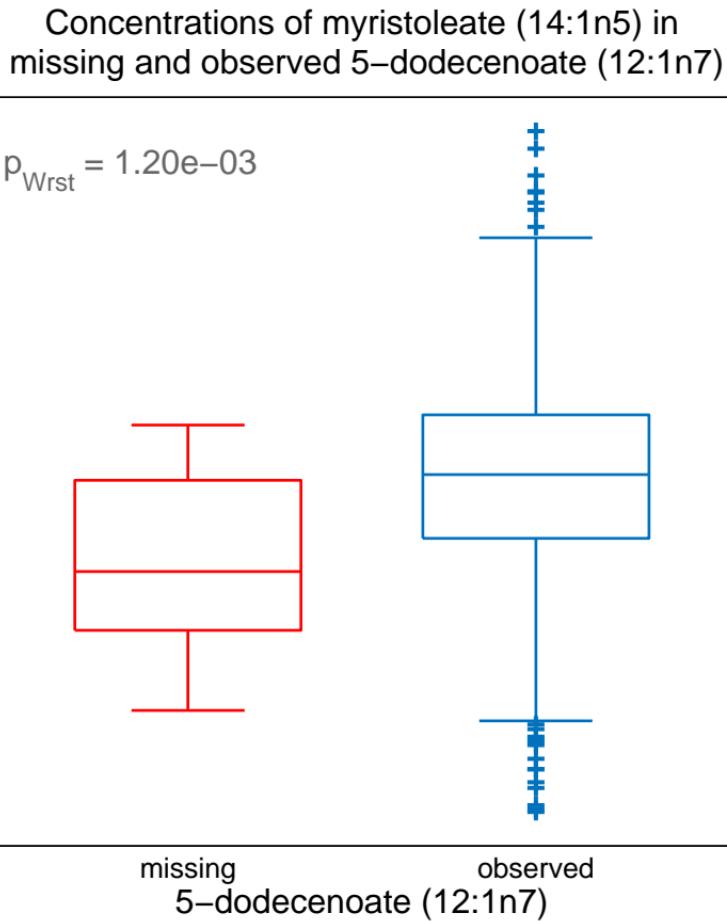
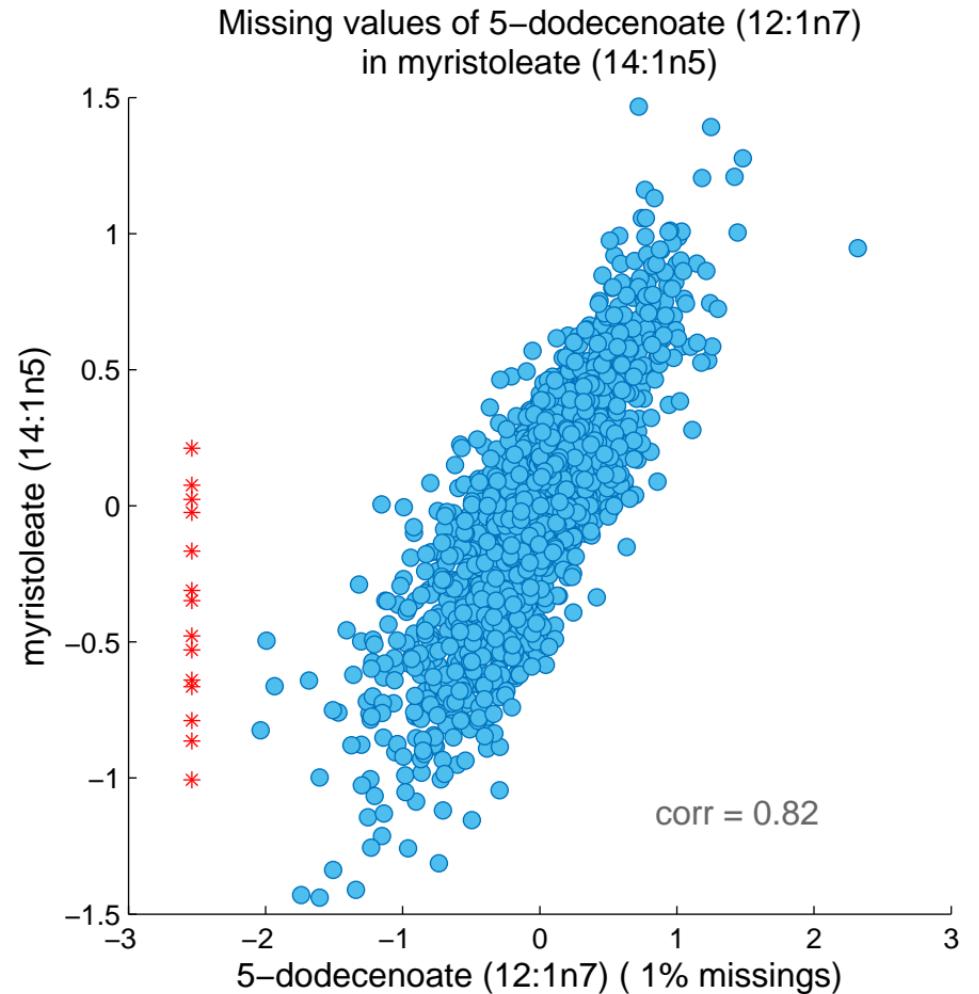




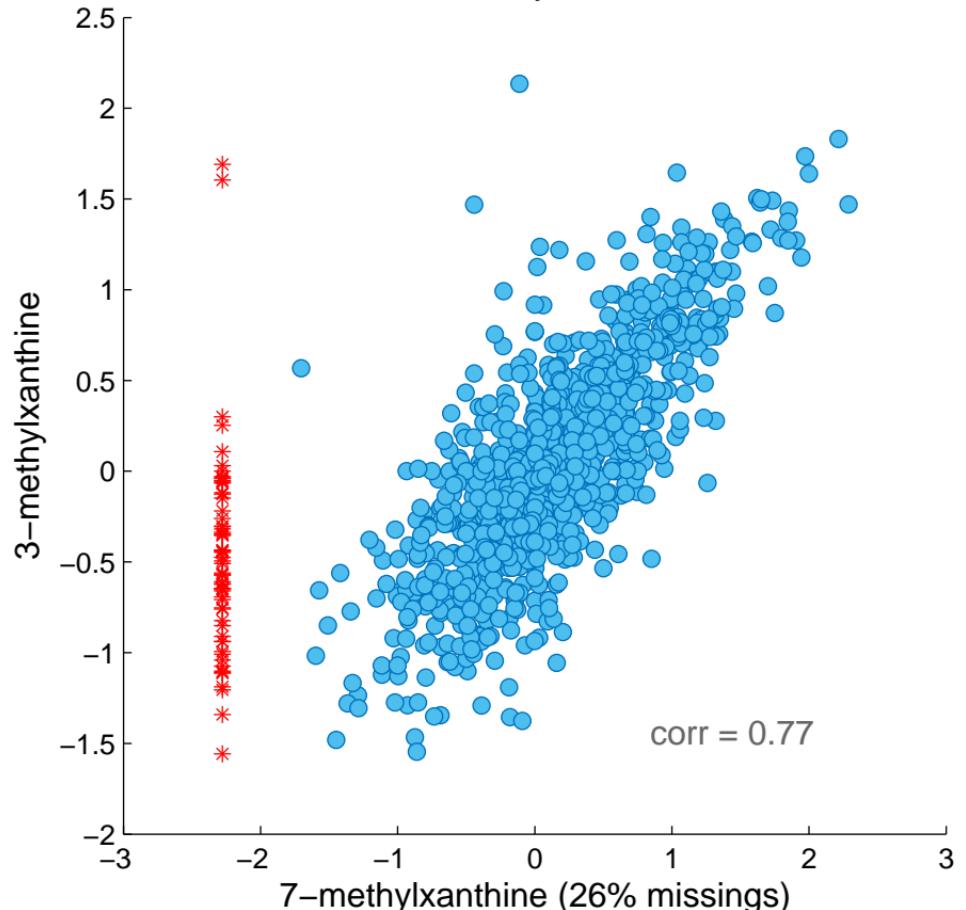




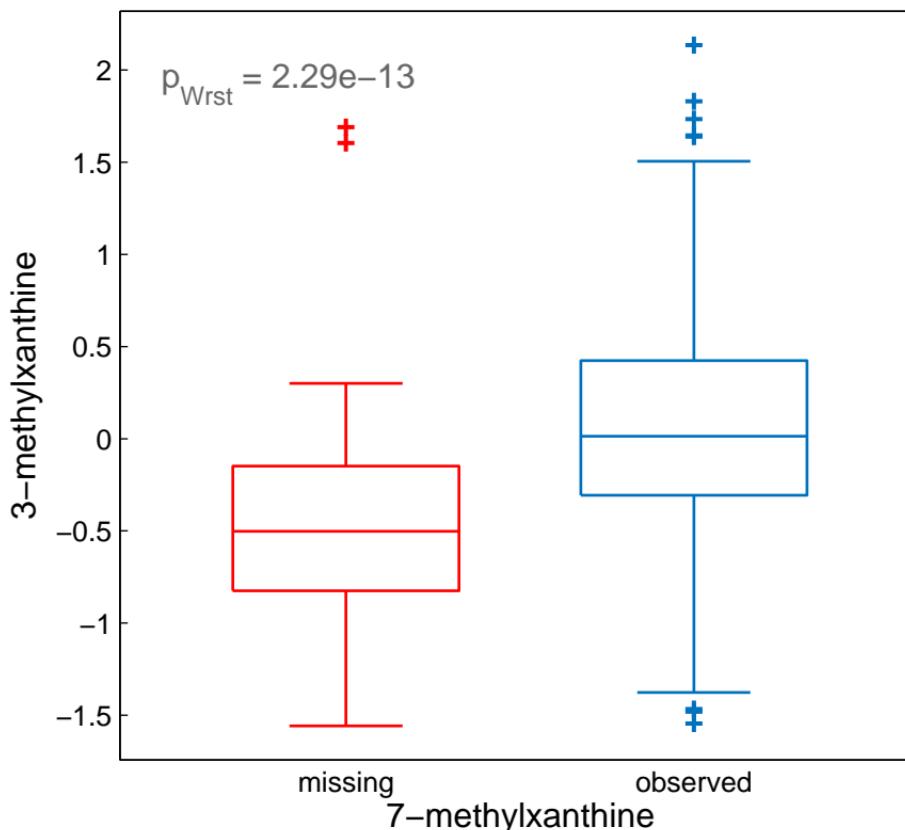


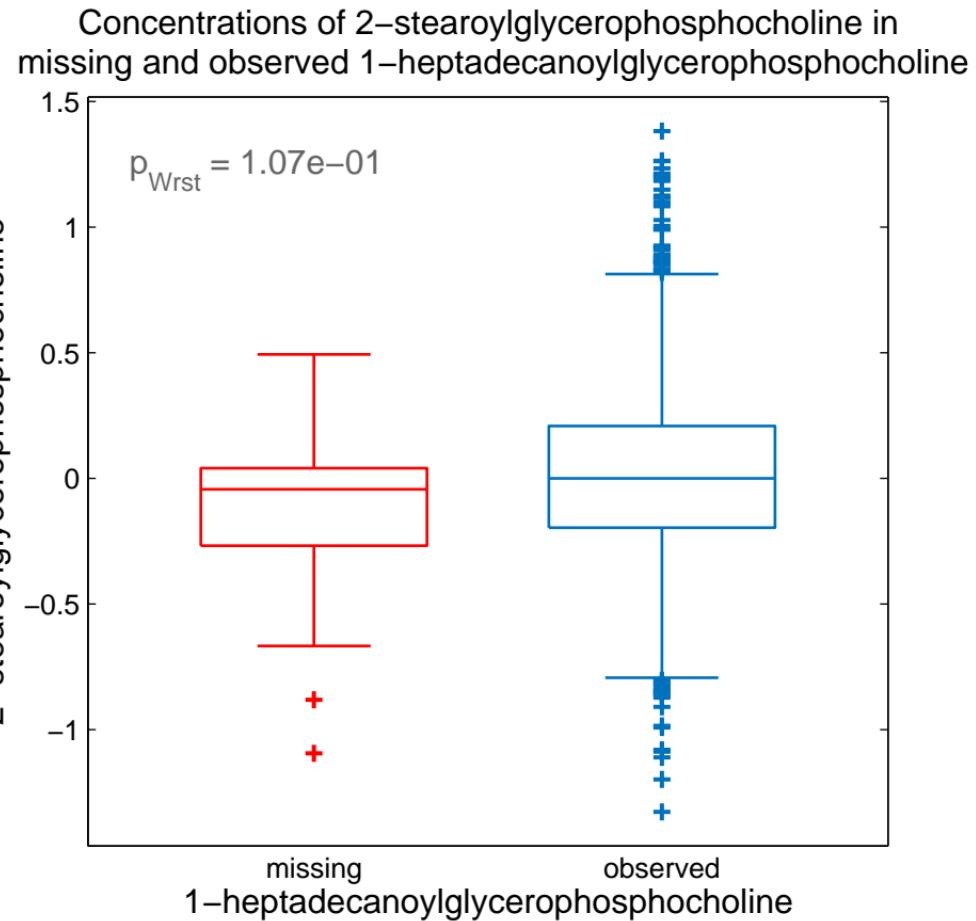
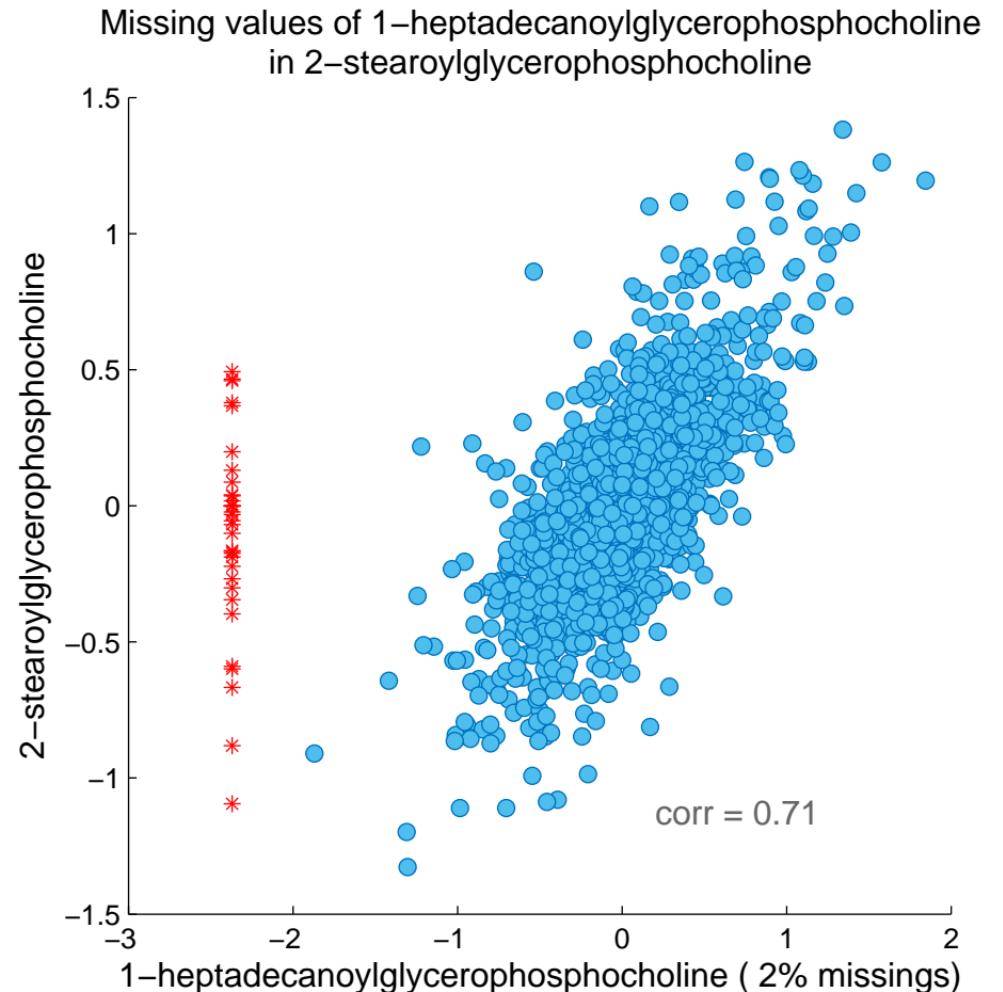


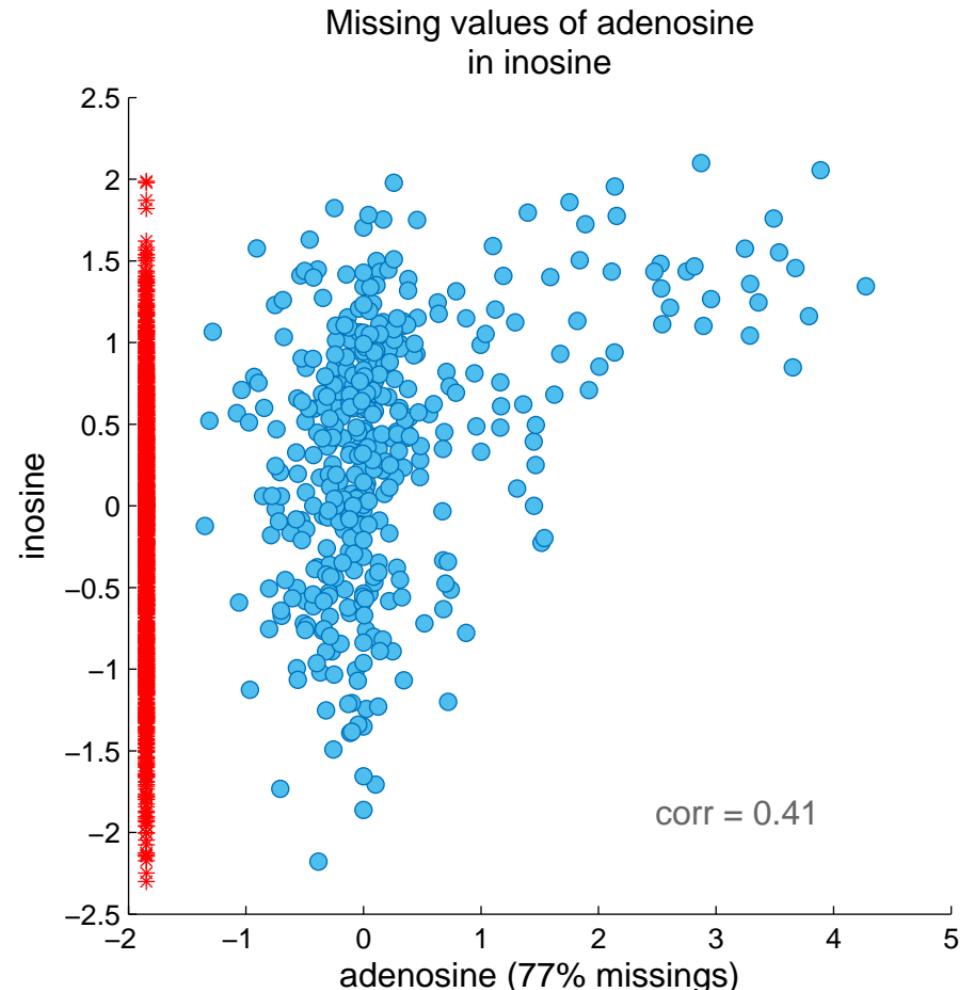
Missing values of 7-methylxanthine
in 3-methylxanthine



Concentrations of 3-methylxanthine in
missing and observed 7-methylxanthine







Concentrations of inosine in missing and observed adenosine

